Université de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :
Institut National Polytechnique de Toulouse (INP Toulouse)

Discipline ou spécialité :

Pathologie, Toxicologie, Génétique et Nutrition

Présentée et soutenue par :
Guillaume Sallé

le : lundi 17 décembre 2012

Titre :

Détection et validation fonctionnelle de régions du génome affectant la résistance aux strongles gastro-intestinaux chez le mouton

---

Ecole doctorale :
Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingénieries (SEVAB)

Unité de recherche :
UMR1225 INRA/ENVT IHAP / UR631 INRA SAGA

Directeur(s) de Thèse :
Philippe Jacquiet
Carole Moreno

Rapporteurs :

Didier Boichard
Mike Stear

Membre(s) du jury :

Stephen C. Bishop
Cédric Neveu Jean-Michel Elsen
Philippe Jacquiet Carole Moreno

Contents

List of Figures

List of Tables

Abstract

This document tackles the problem of finding the genes controlling the susceptibility status of sheep towards gastro-intestinal infection, *Haemonchus contortus* being taken as a model organism.

Part I

Introduction

Les strongles gastro-intestinaux représentent une des contraintes les plus importantes pour l'élevage ovin allaitant. Ces infestations engendrent des pertes économiques majeures, principalement dues aux pertes de production et aux coûts de traitement. Parmi ces strongles gastro-intestinaux, *Haemonchus contortus* est un ver hématophage de la caillette des ruminants dont la pathogénicité et l'ubiquité géographique (régions tropicales, subtropicales et tempérées) lui confère une importance particulière.

La gestion classique des strongles gastro-intestinaux dont *H. contortus* repose sur l'utilisation de molécules anthelminthiques. Depuis les premières avancées de l'industrie pharmaceutique, la mise sur le marché de nouvelles classes d'anthelminthiques a toujours précédé de quelques années l'apparition de vers résistants à ces molécules [350]. Localement ce problème peut s'avérer une menace directe à la survie de l'élevage. C'est le cas par exemple de l'élevage ovin laitier au Pays Basque, qui doit jongler entre un arsenal thérapeutique très réduit et une distribution importante de populations de vers résistantes. Par ailleurs, les attentes sociétales des dernières années envers l'élevage et les denrées d'origines animales sont très nettement marquées par le respect de l'environnement et la réduction d'intrants, contribuant ainsi à la réduction de l'utilisation des médicaments vétérinaires, des anthelminthiques notamment. L'impact environnemental de certaines de ces molécules est par ailleurs avéré [324].

Proposer une solution alternative aux anthelminthiques dans la gestion du parasitisme n'est pas aisé, même si d'importants efforts de recherche ont été fournis [247, 236, 273]. De ces travaux, il apparaît clairement qu'on ne peut se priver complètement de traitements anthelminthiques. En revanche, une utilisation plus raisonnée et optimisée pour ne traiter que les populations d'ovins à risque pourraient contribuer à limiter l'apparition des populations de vers résistants. Cette approche requiert cependant une organisation et une surcharge de travail pour l'éleveur qui tendent à freiner leurs mises en application. D'autres approches visant une meilleure gestion des pâtures présentent également les mêmes contraintes et ne sont pas applicables dans toutes les situations d'élevage. Enfin le développement de vaccins n'a encore jamais abouti à une immunogénicité et une protection durable de l'hôte.

A toutes ces méthodes de gestion alternative s'ajoute la sélection génétique d'ovins plus résistants aux strongles gastro-intestinaux. Cette stratégie ne fait pas encore l'unanimité du monde scientifique étant donné que la théorie de la génétique quantitative prédit que les gènes favorables à la survie sont sélectionnés au cours des générations, aboutissant ainsi à leur fixation. Cependant, de grandes variations de la sensibilité des ovins aux strongles existent.

Les races ovines tropicales sont généralement plus résistantes que les races hautes productrices sélectionnées en milieux tempérés. C'est le cas par exemple des races Martinik Black-belly, plus résistante, et Romane, plus sensible [571, 24]. D'autre part, cette variation est non seulement observable entre races, mais également au sein de populations ovines. Toutes les estimations de la part de la variation observée d'origine génétique ou héritabilité, ont montré que les gènes d'un individu controllaient 30% de la variation observée. Ceci suggère donc qu'il est possible de sélectionner les populations ovines pour augmenter leur résistance génétique moyenne à l'infestation par des strongles gastro-intestinaux.

L'intégration de la résistance aux strongles gastro-intestinaux dans l'indexation des béliers pose cependant nombre de questions sur l'impact d'une telle sélection sur les performances des ovins, sur l'apparition éventuelle de souches de vers contournantes et sur une éventuelle hypersensibilité des moutons sélectionnés à d'autres maladies infectieuses. La mise en oeuvre de la sélection génétique pour la résistance au parasitisme passe donc par une meilleure compréhension des mécanismes sous-jacents à la variation de sensibilité aux vers.

Des études *in vivo* et *in vitro* menées sur des modèles murins ont montré que les infestations par des nématodes gastro-intestinaux entraînent une réaction immunitaire adaptative de type Th2 [17]. Les cytokines sécrétées par les lymphocytes T CD4+ Th2 conduisent au recrutement de mastocytes et d'éosinophiles au site d'infestation ainsi que la production d'anticorps spécifiques [17]. Chez le mouton, ce modèle semble également s'appliquer [288, 490]. De plus, les travaux de comparaison de la réponse immunitaire des races Martinik Black-belly et Romane ont montré une plus forte imprégnation cytokinique de type Th-2 dans la muqeuse abomasale chez la race résistante [507].

Pour préciser le déterminisme génétique des différences observées entre ces deux races, un protocole back-cross de grande échelle a été mis en place à l'INRA de manière à déterminer quelles régions du génome étaient prépondérantes dans l'explication de cette variation [374]. Cinq familles de pères F1 croisés Martinik Black-belly x Romane ont été constituées par croisement en retour avec des brebis Romanes. Au total, 1,000 animaux back-cross ont ainsi été infestés expérimentalement par *H. contortus* à deux reprises pour mesurer leur résistance en primo-infestation ainsi qu'à la ré-infestation. En vue de localiser les régions du génome associées à la résistance, les cinq pères et leurs descendants ont été génotypés pour un nombre important de marqueurs microsatellites. Suite à cette première étape, un consortium international a vu le jour afin de séquencer le génome ovin et de développer une puce à ADN comprenant plus de

50,000 marqueurs génétiques, finalement disponible en 2009.

Profitant de cette avancée technologiue, les travaux de cette thèse ont été réalisés dans le but de préciser l'architecture génétique de la résistance à *H. contortus* en deux étapes. Premièrement, une analyse statistique d'association entre les marqueurs génétiques couvrant le génome et les caractères de résistance mesurés visait à localiser finement les segments chromosomiques expliquant une part non négligeable de la variation génétique. Dans une deuxième étape, une des régions ainsi identifiée a été sélectionnée avant d'être étudiée plus en détails. Cette étape de validation fonctionnelle a été réalisée par des accouplements contrôlés d'animaux back-cross, sélectionnés sur la base des segments chromosomiques dont ils ont hérités des pères F1, afin de produire des animaux porteurs de deux segments à effet favorable ou défavorable sur la résistance à *H. contortus*. Les deux groupes d'agneaux back-cross x back-cross ainsi constitués ont été infestés expérimentalement et des phénotypes fins ont été mesurés pour préciser le rôle de la région génomique d'intérêt.

Ce manuscrit est composé de quatre grandes parties. La première partie est consacrée à une revue bibliographique rappelant la complexité des interactions entre l'hôte et les nématodes gastro-intestinaux et discute la pertinence de la sélection génétique comme méthode de gestion du parasitisme en élevage ovin. La seconde partie présente le matériel et les méthodes utilisés tandis que la troisième expose les principaux résultats obtenus. Une quatrième partie est dédiée à la discussion des résultats obtenus et aux perspectives qui en découlent.

Gastro-intestinal nematodes are a major threat to the meat sheep industry. These infections are responsible for important income losses, mainly due to production losses and treatment costs. Among these gastro-intestinal nematodes, *Haemonchus contortus* is an hematophagous worm living in abomasum of ruminants. Both its pathogenicity and its ubiquity (tropical, sub-tropical and temperate areas) make it the major gastro-intestinal nematode of ruminants.

Classical management of gastro-intestinal nematodes, including *H. contortus*, relies on the use of anthelmintic drugs. Since the first advances of modern pharmaceuticals, the release of new anthelmintic classes has always been followed by the development of resistance to these molecules in worm populations [350]. At a local scale, this issue can directly threaten the sheep breeding industry. This has been the case in the French Pays Basque as dairy sheep breeders face both a limited therapeutical arsenal and a high prevalence of resistant worms. In addition, environmental concerns and the reduction of chemicals in agriculture have shaped

social expectations toward animal production and animal products over the past few years, thus contributing to lower the use of anthelmintics. The environmental impact of some of these molecules has also been confirmed [324].

To propose alternative strategies of worm management is not an easy task, even if great research efforts have been achieved [247, 236, 273]. From this literature, it seems impossible to fully stop the use of anthelmintics. Still, a more integrated and optimized use of anthelmintics to only target the most susceptible sheep subpopulations should contribute to lower the selection of resistant worm populations. However this approach requires a work overload for sheep breeders thus hampering their on-field implementation. Other approaches aiming at a better pasture management also show similar constraints and these are not suitable for every farm. Last, vaccine development has neither achieved a sufficient immunogenicity nor a long-lasting protection of the host yet.

The breeding of more resistant sheep is another alternative strategy for worm control. This strategy still faces adverse controversy in the scientific community that deny the efficacy of such an approach. Indeed, quantitative genetics theory predicts that any gene with a favorable effect on host survival will be passed through generations and should finally be fixed by natural selection [485]. However, great variability in sheep susceptibility has been observed. Tropical sheep breeds usually exhibit more resistant phenotype than their high producing counterparts from temperate area. For instance, the Martinik Black-belly has been demonstrated to be more resistant than the Romane breed [571, 24]. Further, this variation not only occurs between breeds but also within ovine populations. Estimations of the variation explained by the sheep genome called "heritability" found that individual's genes explained 30% of the observed variation. This suggests that selecting sheep populations to increase their average resistance to gastro-intestinal nematodes is feasible.

Introducing resistance to nematodes in rams indices pose a great number of questions among which, the associated impact of this selection on sheep production, the putative selection of worms able to withstand host resistance and a putative hyper-sensitivity of sheep to other infectious diseases. The implementation of genetic selection for resistance to gastro-intestinal nematodes hence requires a better understanding of underlying mechanisms.

In vitro and *in vivo* studies performed on mice demonstrated that infection by gastro-intestinal nematodes elicits an acquired Th-2 biased immune response [17]. The release of cytokines by T CD4+ Th2 lymphocytes further the recruitment of mastocytes and eosinophils

to the infection site and enhance the production of specific antibodies [17]. In sheep, same mechanisms seem to be involved [288, 490]. In addition, the comparison of the immune response between the Martinik Black-belly breed and the Romane breed demonstrated a stronger Th-2 type cytokinic environment in abomasal mucosa of the resistant breed [507].

To better understand the genetic determinism of these differences, a high-scale back-cross design has been implemented to determine which genomic regions explained most of the observed variation [374]. Five F1 sires generated by crossing Martinik Black-belly and Romane have been back-crossed to Romane ewes. The 1,000 back-cross progenies were experimentally infected twice by *H. contortus*, in order to measure their resistance in both primary and secondary challenge. To map regions of the genome associated to resistance, the five F1 sires and their progenies were genotyped for a large panel of microsatellite markers. Following this first step, an international consortium was built to sequence the ovine genome before developing a DNA SNP chip of more than 50,000 genetic markers that was subsequently released in 2009.

Benefiting this technological advance, this two-step PhD project aimed at mining the genetic architecture of resistance to *H. contortus*. Firstly, regions of the genome significantly associated to resistance traits and explaining a non negligible part of the genetic variation were mapped by statistical analyses. Secondly, one of the identified regions was selected for an in-depth characterization. This functional validation step was achieved by marker-assisted matings of back-cross sheep, selected based on the chromosome segment they inherited from their F1 sires, in order to produce sheep carrying two favorable/unfavorable alleles. Two groups of back-cross x back-cross lambs were constituted and experimentally challenged, before being followed up and intensively phenotyped for a wide range of traits hence providing additional insights on the role of the region under study.

This manuscript has been divided in four parts. The first chapter has been dedicated to a literature review recalling the complexity of the sheep-nematode interactions and discussing the implementation of genetic selection as a control strategy. A second part introduces materials and methods used in this project while the third part reports main results. A last part discusses the results and the associated perspectives.

Part II

Literature review

A review of current knowledge in the field of genetic resistance of sheep to GIN is provided. It proposes the why and how of using genetics for increasing resistance of sheep populations to GIN infection.

This review recalls the existing interplay between GIN and sheep from epidemiological and immunological perspective.

The opportunities for using genetics are then considered in the same way as what has been proposed by Davies *et al.* [107], *i.e.* considering seven different socio-economical parameters through extensive bibliographic reviewing.

Remaining parts of this literature review are devoted to the different strategies of implementing genetic solutions for controlling GIN in sheep breeding. Classical selection is considered followed by a complete description of the available molecular tools and their applications, methodologies related to the detection of genes of interest being tackled in the last part.

Chapter 1

The complex interplay between sheep and gastro-intestinal nematodes

This part aims at recalling basic knowledge about major nematodes species.

1.1 A brief review of sheep GIN

Gastro-intestinal strongyloses are due to the presence and the development of *Strongylids* in the abomasum, small or large intestine after the ingestion and/or the transcutaneous penetration of infective larvae [306]. Even if more frequent in warm areas, these infections are ubiquitous. This is particularly well illustrated by the *Haemonchus contortus* species that can be found in northern Canada regions as well as in humid tropical countries. The two "ecotypes" hence exhibit different behaviour: most of 3rd stage larvae are killed by frost but all the 4th stage larvae enter hypobiosis stage during artic winter whereas almost no hypobiosis can be observed in temperate climate and high survival rate is observed in 3rd stage larvae both in winter and summer. The classification is as follows:

- Class: Nematodes
- Under-class: Secernentea
- Order: Strongylidea
- Super-family: Trichostrongyloidea

Within this super-family two families have a particularly important weight in veterinary medicine:

- Trichostrongylidae This family contains the three most abundant and economically relevant nematodes species in sheep breeding, i.e. *Haemonchus contortus*, *Trichostrongylus colubriformis*, *Teladorsagia circumcincta* [395].
- Molineidae among which *Nematodirus* is the most pathogenic for sheep.

Other relevant genera, *i.e. Oesophagostomum*, *Chabertia*, *Bunostomum*, belong to the Strongyloidea super-family.

1.2 Macro-interactions between sheep and nematodes: sheep GIN lifecycle and epidemiology of strongyloses

1.2.1 Lifecycle of sheep GIN and associated pathology

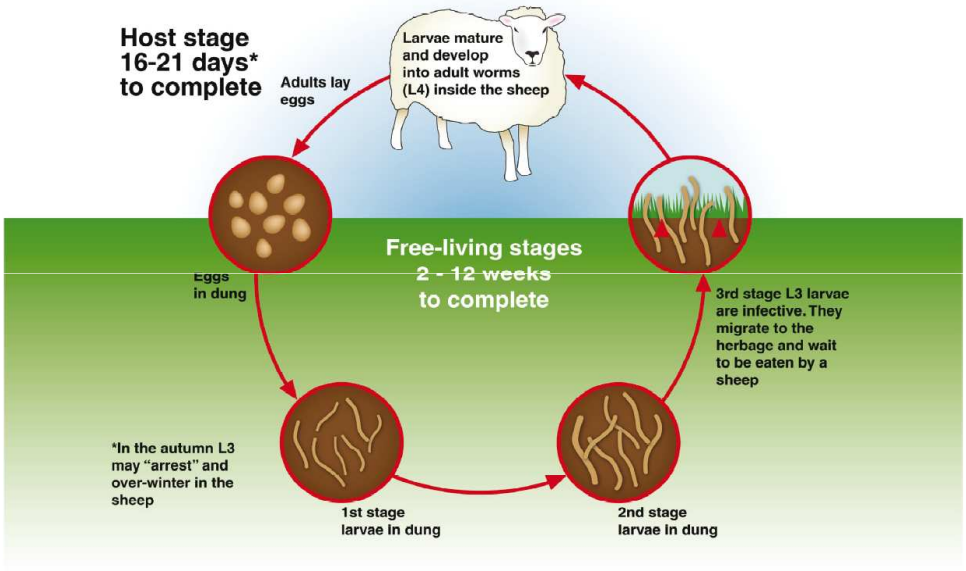
The GIN lifecycle is direct (without an intermediate host) and can be split in two distinct parts, within and outside the host as illustrated on figure 1.1.

The worm life cycle begins with the ingestion of third stage larvae (L3) present on pasture. Larvae subsequently evolve into the following stage (L4) that can arrest development and overwinter within the host hence surviving the cold winter temperatures [95, 396, 149]. This is particularly true for *Haemonchus*, *Teladorsagia* and *Cooperia*. After a short migration through the digestive tract, larvae continue their evolution until reaching their final destination, *i.e.* abomasum in the case of *H.contortus* and *T. circumcincta* or a bit further to the gut for *Trichostrongylus spp.* or *Nematodirus spp.*. After mating, female worms will shed eggs that will be subsequently excreted on pasture and will finish the loop by evolving into an L3 larvae.

Female worm fecundity varies according to the parasite species [498]. Two different profiles can be proposed:

- Adults with short life but high fertility, like *Haemonchus contortus* females that usually shed more than 6500 eggs/day
- Life span longer for adults but a reduce fertility of females, like *Trichostrongylus colubriformis* that produce around 260 eggs/female/day

Figure 1.1: Gastro-intestinal nematode life cycle



Reproduced from Abbott *et al.* (2007) [2]

The prepatent period, that represents the time frame between the ingestion of L3 larvae and the final egg sheddings, usually lasts two weeks with some variations according to the parasite species, the climate or the host immune reponse.

Early larval stages present on pastures feed on bacteria present in feces while L3 do not feed and migrate to the grass. They are isolated from the environment by the L2 stage cuticule that both protects them and prevents them from feeding [498]. Within the host, parasites feed on mucosal fluids, digestive products from their hosts and cellular pieces. The *Haemonchus* genus shows some differences as it sucks blood from its host as soon as the L4 stage.

1.2.2 Epidemiology and ecology of sheep GIN

Lambings usually occur at the beginning of the spring and both ewes and lambs are put on pastures a few weeks after [498]. Pasture contamination in the spring comes both from overwintered infective L3 on pasture, and from nematode eggs shed by recently-lambed ewes.

Overwintering of L3 larvae is determined by the capability of the eggs to evolve to infective stage, an evolution conditioned by temperature and moisture [395]. To this regard, eggs of *T.*

Table 1.1: Ecology of nematode immature stages from egg to L3 stage (reproduced from O'Connor, 2006 [395]

Nematode species	**Stages**			
	Unembryonated egg	*Embryonated egg*	*Pre-infective larvae*	*Infective larvae*
H.contortus	Highly susceptible to cold and dessication. High mortality below 10C	Susceptible to cold and dessication. Negligible hatching below 10C. Low hatching rates in absence of moisture	Highly susceptible to cold and dessication	Optimum survival in warm, moist weather. Poor survival in cool or warm dry weather and sub-freezing winters
T.colubriformis	Intermediate susceptibility to cold and dessication. High mortality below 5C	Intermediate susceptibility to cold. Low susceptibility to dessication	Susceptible to cold. High mortality below 5C. Susceptible to dessication	Optimum survival in cool or warm moist weather. Poor survival over sub-freezing winters
T.circumcincta	Low susceptibility to cold. High egg viability at 0-10C. Intermediate susceptibility to dessication	Low susceptibility to cold and dessication. Hatching below 5C.	Intermediate susceptibility to cold. Susceptibility to dessication	Optimum survival in cool, moist weather and sub-freezing winters. Poor survival in warm, dry weather.

circumcincta can reach the infective stage at colder temperatures than *T.colubriformis* whereas *H.contortus* is the most affected by cold temperatures (reviewed by O'Connor [395]). Eggs of this genus are also the most affected by dessication [395]. Their relative sensitivity to environmental conditions is partly compensated by their ability to migrate to favorable micro-environments and by their sheath, that is likely to protect them against dessication [395]. Experiments at constant temperature of $-10°C$ showed that *H.contortus* L3 survived only 24 hours, whereas *T.circumcincta* and *T.colubriformis* stayed alive for three months and eight days respectively [395]. Majors factors affecting stages from egg to infective larvae are summarized in table 1.1.

The increase in egg excretion of recently-lambed ewes is well described and known as the "periparturient rise". It can begin up to four weeks before lambing and lasts until eight weeks after [498]. It has been suggested that a variation of prolactin plasmatic concentration could induce a hole in the ewe's immunity. Another hypotheses propose that nutrition is the most important factor : the increase of ewes' nutritional requirements would not be completed as the feeding capacity is reduced during pregnancy. This reduction of intake concomitant to big increase of the needs could result in a hole in immunity. The periparturient rise could be due to the followings [498]:

- Awakening of the hypobiotic L4
- Increase of the parasitic load due to the increased egg output
- Increase of the female worms' fertility

Following the arrival on pastures, naive lambs will feed on infective L3. These larvae will be able to multiply easily as lambs' immune system will not be mature until three months of age. This will increase parasitic load on the pasture until reaching a summer peak with clinical signs shown by lambs and even mortality.

Lastly, hypo-biotic L4 stages will overwinter within housed animals and the cycle will start again on the following spring.

From this rather schematic description of parasitic epidemiology, it appears that some individuals are particularly at risk (naïve lambs, periparturient ewes). This difference in susceptibility is one reason of the aggregation of parasites in their host population, meaning that a large number of hosts harbor a few parasites, while a few hosts are heavily infected [169]. Two additional aspects also condition this over-dispersion that are the spatial dispersion of infective stages and/or environmental and demographic stochasticity [169].

1.3 Micro-interactions between sheep and nematodes: immune response and immunoregulation

Trichostrongylids like *H. contortus* undergo several major modification in the host to evolve from larval stage to the adult stage. These steps include ex-sheathing associated to the complete loss of their glycoprotein rich cuticle or production of excretory/secretory (ES) products that all represent potential antigens susceptible to elicit an immune response (see figures 1.2 and 1.4) [223, 279]. Complete understanding of the host/nematode interactions are far from complete [414, 17]. This section aims at summarizing current knowledge about immunological features during nematode infection by considering a kinetic framework from the first step of the host response until the mounting of an effective memory response. Insights from both murine and ovine models are provided. Due to the interacting aspects of these biological phenomenon, a subsection will be dedicated to the immuno-regulation performed by nematodes.

1.3.1 Keep out property: physical and biochemical barriers

During GIN infection, the mucosal surface of the gastrointestinal tract is the interface between the host and its pathogen. To avoid invasion, this physical barrier can act through mechanical means like fluid movement and peristalsis associated with biochemical defences [115]. For instance, low pH and digestion enzymes make the sheep abomasum a particularly hostile environment for pathogens.

In addition, epithelial cells secrete lectins that are proteins that bind to glycosylated molecules [528]. Lectin molecules are ubiquitous and have been found on both mammalian and microbial pathogens. In sheep, galectin-15 has been found in sheep trickled with *H.contortus* but not in normal gastrointestinal tract [136]. This result was subsequently repeated for *T. circumcincta*, either by proteomic study [412] or expression study [438, 277]. In this latter study, galectin-15 was significantly up-regulated as soon as day 3 post challenge and peaked between the 5th and 7th day of challenge [438]. Furthermore expression studies performed after *T. circumcincta* infection in sheep demonstrated the up-regulation of the intelectin-2 protein in the early step of infection [166, 277]. It is thought that these proteins bind both to mucins of the mucus and to the parasite surface, thus increasing both mucus viscosity and its adherence to parasite [115, 528]. Such entrapment would then facilitate nematode expulsion.

Goblet cells of the epithelial mucosa are also the source of trefoil factor-3 (TFF3) that seems

to have a protective role against damage [21]. The *TFF3* gene was found up-regulated at local site of infection from day 3 to 21 after *H.contortus* or *T. colubriformis* challenge [359]. The same kind of findings were reported by Knight *et al.* who reported up-regulation of *TFF3* in immune sheep challenged with *T. circumcincta* [277]. In addition, the calcium-activated chloride channel 1 (*CLCA1*), that is a mucous cell associated transcript, was one of the most up-regulated transcripts seen in immune sheep [277] and its role has been also reported in mice infected by intestine-dwelling nematodes [21].

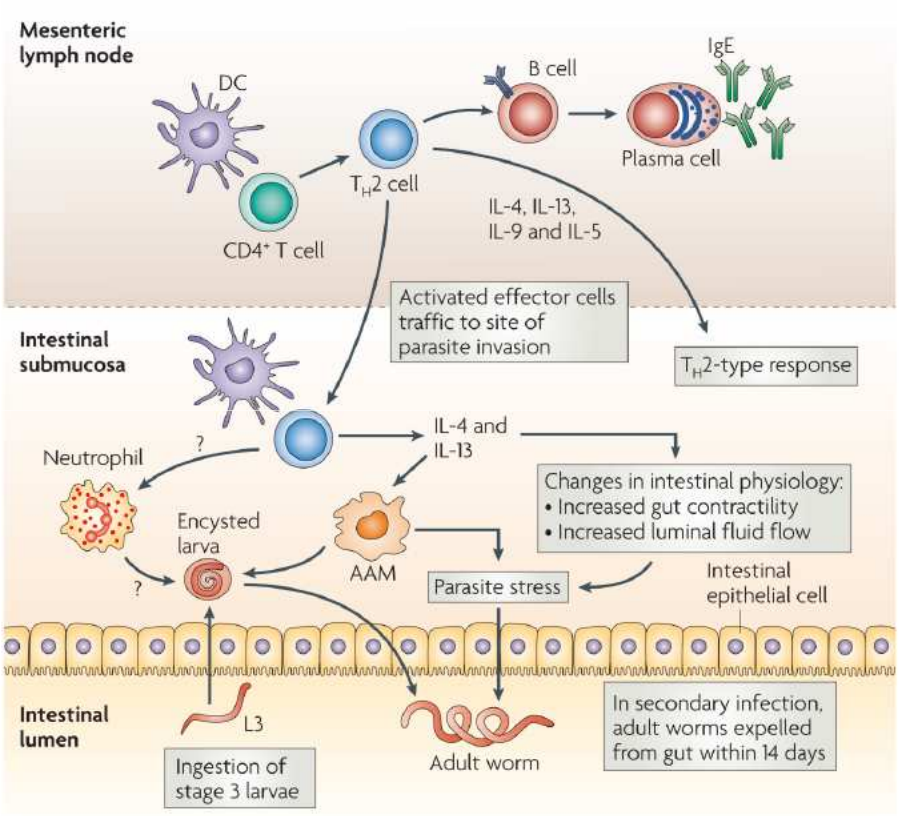
Furthermore, there is evidence that proteins usually associated with the regulation of food intake may play a role in nematode elimination. This has been shown in mice where expulsion of *Trichinella spiralis* has been associated to the satiety factor cholecystokinin [21]. In sheep, ghrelin expression explained more than 40% of the FEC variation observed after both *H.contortus* and *T. colubriformis* [244]. This protein is secreted by gastrointestinal mucosa [283] and has been known to increase food intake as well as having influence on inflammatory states [244]. However a study performed in calves infected with *O. ostertagi* did not find any evidence of this protein being associated with nematode infection [162].

As outlined by Artis & Grencis (2008) [21], epithelial cells do not only exert a passive mechanical role to nematodes but actively contribute to their elimination, either in mice models or in sheep [277, 359, 245]. They also exhibit some sensing capacities that are developed in the next subsection.

1.3.2 Breaking into the host triggers the innate immune system

Beyond barriers developed to limit pathogen invasion, a network of watchmen constitute the first line of defence known as the "innate" immune system (see figure 1.3). These front line guardians are present in the tissues so that they can respond very quickly to any pathogen. In addition to these cells, soluble factors known as complement that are also ubiquitous and constitutive of the first step of the immune response. All these effectors are fundamental to the mounting of a more pathogen-specific response but the complete delineation of the initiation of such response remains under investigation [115, 21]. Identifying pathogen requires host immune cells to have some way of fingerprinting pathogen as well as differentiating them from the host cells itself. A brief reminder on self and pathogen recognition is provided as well as major hints about the innate immune system in nematode infection.

Figure 1.2: Schematic representation of a protective Th2 immune response against helminths in mice



Reproduced from Anthony *et al.* (2007) [17]

1.3.2.1 The concept of self has evolved since 1959 [343]

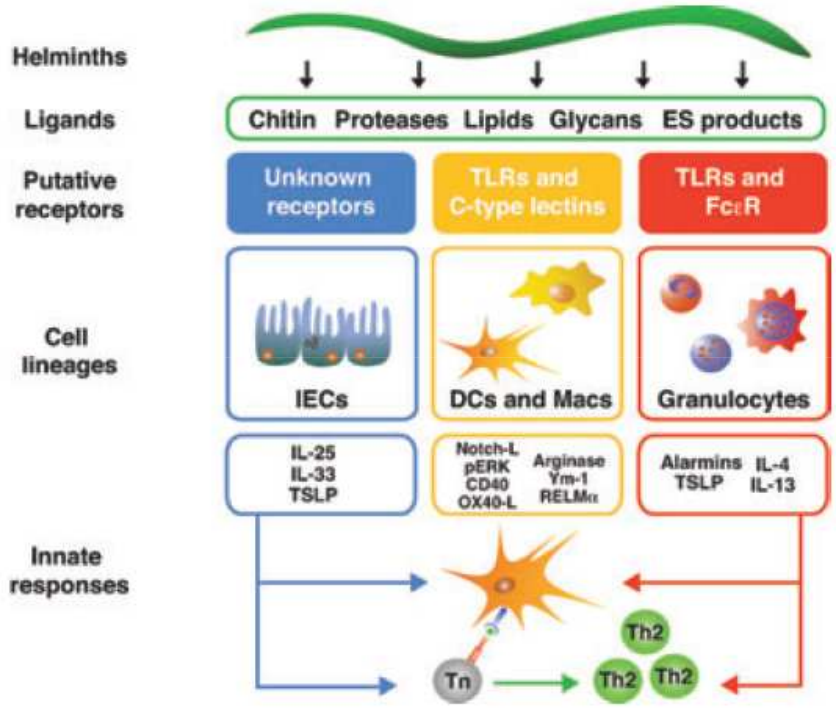
The original model of self and non-self recognition was proposed by Burnet in 1959 [74]. He stated that lymphocytes express receptors specific for an exogenous entity, that recognition of this entity initiates the immune response and that self-reactive lymphocytes are eliminated to avoid auto-immunity [74].

Subsequent findings in immunology were not in perfect acquaintance with this model and Matzinger recently proposed a "danger" model in which APCs are activated by alarm signal emitted by distressed cells, *e.g.* cells exposed to pathogens [343]. This model better explains why the immune system reacts to necrotic cells but not to senescent dying cells.

It seems that the growing interest for pathogen recognition mechanisms have produced results in agreement with this new model but very few studies have applied this concept to nematode infection [115].

1.3.2.2 Fingerprinting of the intruder (figure 1.3) [414, 115, 147, 322]

Figure 1.3: Innate immune cell recognition and response to helminth-derived products



Reproduced from Perrigoue *et al.* (2008) [414]

Gastro-intestinal nematodes express a wide range of proteins, produce many Excretory-

Secretory (ES) products and exhibit a large surface in close interaction with their hosts, *i.e.* cuticle and gut [223]. All these molecules thus represent potential antigens that can be identified by the host as "non-self" entity hence launching an immune response to expel invaders [223, 388]. These somatic and ES antigens are known as "natural" antigens [388]. They are in contact with the host immune system and are to be opposed to "hidden" antigen (see section 1.4) [388].

Among natural antigens, several proteins and glycoproteins from *H. contortus* [388] and *T. colubriformis* [142, 143] demonstrated immunizing properties [212]. Among identified proteins are the Hc-sL3 antigen which is specifically expressed on ex-sheathed *H. contortus* L3 larvae [248, 388], and the "CarLA" (for Carbohydrate Larval Antigen) that is expressed on the surface of strongylid nematodes infective larvae and that has been associated to *T. colubriformis* rapid elimination in sheep [212, 211, 210].

Some natural antigens have been identified but they do not confer sufficient protection for young naive animals against nematode challenge [388], as recently exemplified by the Tc-SAA-1, a *T. circumcincta* specific antigen [393]. Additional experiments are also needed to test the immunogenic properties of the identified peptides.

1.3.2.3 Nematodes activate the complement through the alternative way

Complement is composed of soluble mediators secreted by the liver and with an ubiquitous dispersion in the host organism, either in blood or tissues [115].

The importance of complement during GIN infection is poorly understood and few data is available [115, 177]. It is thought to be a first-line defense that is rapidly overcome by the parasite [177].

In sheep, it has been demonstrated that complement binding to eosinophils facilitates their action [427] and some complement factors were up-regulated after *T. circumcincta* challenge [277].

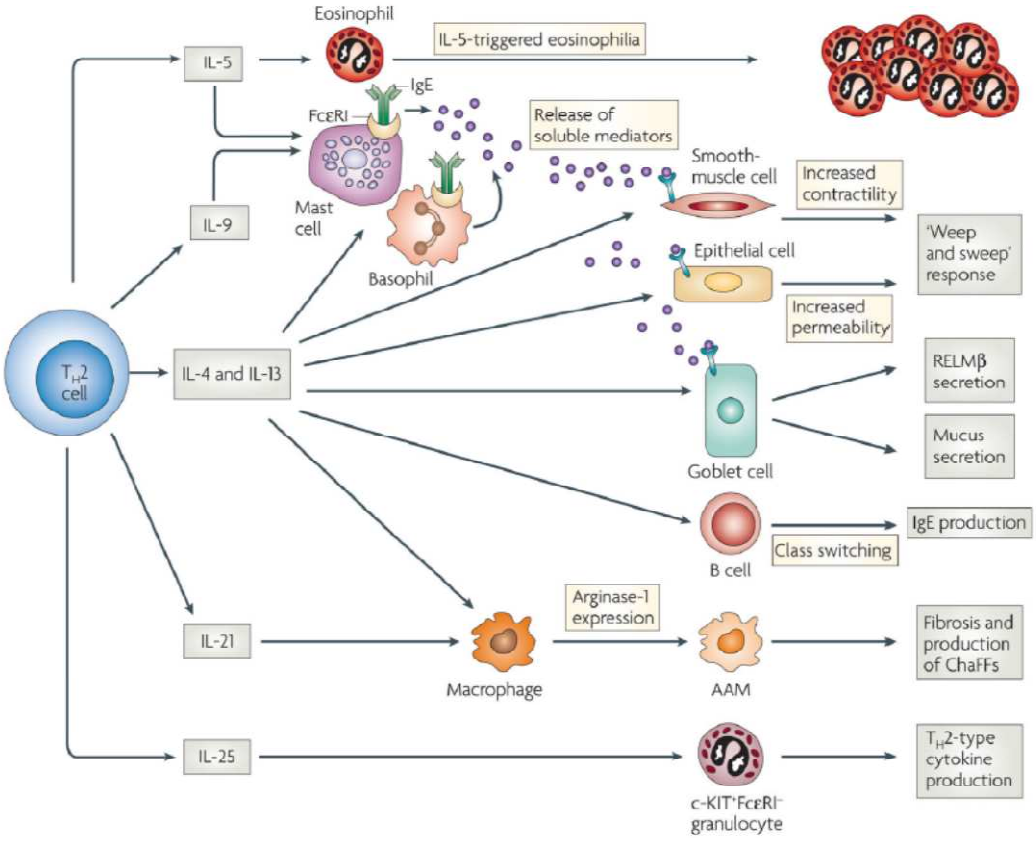
The different interactions between the innate immune system and the GIN antigens is responsible for a cascade of activation that contribute to mounting of a pathogen-specific immune response that we detail further, as illustrated on figure 1.3.

1.3.3 Experts step in for an efficient acquired response

Even if the innate immune system represents a solid barrier against intruders, it does not mount a specific response and acts in an "innate" fashion. On the opposite, the acquired immune response

relies on a specific antigen recognition by APCs (dendritic cells, macrophages, B lymphocytes) that furthers the selection and activation of specific T lymphocytes (see figure 1.4). Two types of T lymphocytes can be distinguished by their cluster of differentiation (CD), *i.e.* CD4+ and CD8+ [498]. T cells are equipped with receptors (TcR) that specifically recognizes epitopes presented by the Major Histocompatibility Complex (MHC) exhibited on cells surface.

Figure 1.4: Effector cells in the Th2 immune response against helminth



Reproduced from Anthonly *et al.* (2007) [17]

CD8+ T cells are the most frequent intra-epithelial lymphocyte type in ruminants [513]. They target MHC class I molecules that bind intracellular epitopes. These cells are thus more dedicated to the response against intra-cellular pathogens. On the contrary, CD4+ T cells are dedicated to the recognition of antigens presented by MHC class II molecules. CD4+ T cells are critical players in the immune response. Once activated CD4+ T cells differentiate into T helper (Th) cells that polarize the immune response through the release of specific cytokines [371]. In murine model, two subsets of Th, *i.e.* Th1 and Th-2, have been identified with various

associated cytokines (see table 1.2).

Table 1.2: Main cytokines and associated effects [498]

	Th-1	**Th-2**
Main cytokines	IL-2, IFN-γ	IL-4, IL-13, IL-5, IL-6, IL-10
Stimulated cells	Cytotoxic T cells NK cells, Macrophages	Memory B cells, Plasmocytes, Eosinophils, mastocytes
Inhibited cells	Th-2 cells	Th-1 cells

Key: IL : interleukin, IFN: interferon, NK: Natural Killer

This murine Th cells compartmentalization seems to apply in ruminants [184, 288]. Murine insights have shown that a Th-2 cytokine environment was associated to resistance, but this does not hold for every host-nematode combination [371].

In sheep, a depletion in CD4+ T cells in Gulf Coast Native animals has been demonstrated to confer susceptibility to *H. contortus* infection [411]. Kinetics of the immune response showed that the increase of CD4+ in the draining lymph node occurred as soon as day 3 post-infection [30]. In addition, the comparison of genetically resistant Merino lambs and random bred lambs showed that resistance was associated to higher IL-5 and immunoglobulin G1 (IgG1) and IgE concentrations, hence supporting the Th-1/Th-2 balance in ruminants [184]. Subsequently, Lacroux *et al.* reported an unequivocal Th-2 biased response in infected Romane lambs in comparison to naive individuals [288]. A comparison between susceptible Romane and resistant Martinik Black-belly lambs suggested that this response was higher and more sustained in the resistant breed [507].

During the course of the Th-2 response, B cells are activated and produce various Ig classes that either neutralize the worm or stimulate other cells, *i.e.* IgE and mast cells [17]. It is thought that IgM play a minor role in contrast with their IgG counterpart. In addition, variation in IgA, that are locally secreted in the gut lumen, explains 40% of the variation of female *T. circumcincta* fertility [304]. B-cells have been shown to contribute to the expansion and maturation of protective Th-2 cells through IL-2 secretion in a murine model [562].

1.3.4 The acquired immune response also recruits innate effectors

During Th-2 polarization, IL-4 induces the production of IgE by B cells. IgE subsequently triggers mast cell degranulation. Meanwhile the increase of IL-5 concentration recruits eosinophils at the site of infection. These two phenomena contribute to the creation of an unfavorable

inflammatory environment [17].

1.3.4.1 Eosinophils

The major role of eosinophils in nematode expulsion have long been recognized [355, 443, 541, 56]. They contain granules filled with diverse cationic proteins with cytotoxic and pro-inflammatory properties [443]. They also express TLR receptors and can act as APC [443]. Both these two functions make them potent effectors of anti-GIN response.

Experimental evidences of this role have been gathered from both murine and sheep infection model. In mice studies, it has been shown that (reviewed by Rothenberg [443]:

- eosinophils can mediate helminth killing
- there is an intense infiltration of eosinophils in the vicinity of worms
- IL5, that spurs the eosinophil maturation, plays a protective role in various nematode infection in mice

In sheep, eosinophil number increases in abomasal mucosa five days after *H. contortus* infection, the recruitment being quicker and more directed towards invading larvae in immune sheep relative to naive sheep [30, 356]. In addition, it has also been demonstrated that they are able to inhibit larval motility [427, 32]. Additional *in vitro* exposure of infective *H. contortus* larvae to eosinophils was proven to reduce their establishment potential *in vivo* [506].

While comparing hypersensitised sheep with naive and immune, Kemp *et al.* showed no difference between groups in eosinophils infiltration [268]. Still, using an *ex vivo* abomasal tissue model, kemp *et al.* showed that another galectin, *i.e.* galectin-14, gradually increased with the infection dose of *H.contortus* [268, 572]. This finding suggests that eosinophils might be under continuous recruitment and stimulation hence leading to an accumulation of galectin-14 [268].

In an experimental work by Terefe *et al.* [508], a putative difference in functional properties of eosinophils between two breeds with different resistance status to *H. contortus* was tested. The authors demonstrated a higher infiltration in the resistant Black Belly breed but could not find any evidence in their ability to kill larvae.

The outlined array of data, place eosinophils as key players in the immune response to helminths. Still their activation seems to be under the control of the environment and deserve further experiments [356, 508].

1.3.4.2 Mast cells

Like eosinophils, these cells are filled with many granules but with different content, *i.e* histamin, heparin, proteases. By the end of an infection, mast cells undergo changes to globule leukocytes [371]. Mice models of mast cell depletion, either mast-cell deficient mice or induced depletion through the use of specific antibody, showed that worm expulsion was delayed [21]. In sheep sensitized by repeated infection with *H. contortus* larvae, a rapid rejection of incoming larvae has been observed [31]. This process called "immune exclusion" is associated with strong level of mast cells and globule leukocytes [371, 268, 277].

The release of their granules mediated by Immunoglobulin E (IgE) is responsible for a so-called type I hyper-sensitivity reaction that is involved in the rapid rejection of nematodes [115, 171]. Among the released components, mucosal specific mast cell proteases seem to be particularly at stake during *T. spiralis* infection in mice [21]. It generates inflammation that creates an unfavorable environment for the parasite development as well as an increase of mucosal permeability that facilitates the expulsion of worms in association with an enhanced smooth muscle contraction [346]. It finally results in a "weep and sweep" response [17].

However some worm expulsion can occur without mastocytosis and mastocytosis is not always synonymous of worm rejection [115]. In addition, the role of IgE in nematode response remains controversial [115, 21].

Mast cells along with eosinophils are two of the most important innate immune cells for nematode rejection. It is probable that these two cell types work together in close relationship. Interestingly, mast-cell released mediators are responsible for tissue injury that is known to stimulate eosinophils infiltration and some findings demonstrated that eosinophils carried histamine receptors capable of mediating their activation [115].

1.3.4.3 Other cellular sentinels

In addition to eosinophils, neutrophils and basophils are additional granulocytes involved in the innate immune system. Even if neutrophils and natural killer cells are not typically associated to nematode infection [115], basophils can participate in releasing mediators that contribute to nematode expulsion [540]. However these rare cells have been poorly characterized and their interaction with helminths have been tested in mice model only [540].

From recent studies, it seems that macrophages are not only dedicated to responses to viruses and bacteria [17]. These immune system scavengers can be alternatively activated by IL-4 and

IL-13 cytokines, thus resulting in the lack of upregulation of their inducible nitric oxyde synthase [115]. The alternatively activated macrophages show three different functions that are regulation of the immune response, wound healing and resistance to parasite infection [17].

1.3.5 Applying brakes and maintaining balance [332, ?]

During the infection, the Th-2 response produces some cytokines that recruit effectors which themselves contribute to the amplification of the cytokinic Th-2 environment. The resulting high inflammatory state resulting from this response needs to be controlled to avoid any deleterious effects for the host.

In mice models of nematode infection, the regulation of the immune response has been shown to be under the control of a particular subset of T cells called T-reg, either being naturally occuring or induced. Still, their precise characteristics remain elusive. It is known that removal of Treg in nematode infection in mice can further worm death but also inhibits the priming of Th-2 response.

Alternatively activated macrophages (AAMs) were firstly thought to promote pathology during chronic infection. However, insights from mice model demonstrated that a resistin-like protein secreted by AAMs was required in the expulsion of *H. polygyrus*. Additional findings from filarial infection suggested that AAMs could play a role in the direct killing of parasites. A study by Anthony *et al.* in mice infected by *Heligmosomoides polygyrus* also demonstrated that AAMs were important effector of the protective memory response.

1.3.6 An adapted response against a moving target

Through evolutionary process, GIN have evolved to establish in their hosts in highly unfavorable environments, capable of mounting a dedicated immune response to expel them out. Therefore it is no wonder that GIN also dispose of an evasion arsenal [226].

Among helminths, a broad range of immunomodulatory molecules have been identified with various properties [226]:

- Production of antioxidants and proteases that inhibit innate cells function
- Cytokines and lectins that disturb recognition and signalling
- Acetylcholinesterases thought to inhibit the fluid increase in the infected gut

In ruminants, identified immunomodulatory ES products seem to be dominated by galectins like in *H. contortus* and other venom allergen [226]. It is has been demonstrated that the *H. contortus* secreted galectin was capable of modulating the eosinophil migration *in vitro* [516]. The precise function of this protein in the modulation of the host response remains to be elucidated. However, it seems that the downregulation of the host response is mostly due to adult stages whereas larval stages are associated to vigorous inflammatory response [223].

Besides the ability of helminths to dampen their host immune response has been exploited for auto-immune disorders therapy purposes [146]. For instance, *Trichuris suis* therapy has been successfull in treating inflammatory bowel disease in humans [493] and some findings indicate that people infected with helminths showed less chronic allergy [21, 146]. However other findings were contradictory and no clear picture can be drawn so far [146]. In addition, side effects of helminthic products in humans is unknown and potential dangerous drawbacks from such therapy like anaphylactic reaction and cross-reactivity to other allergen cannot be ruled out [146].

1.3.7 Consequences of the immune response

The immune response can result in the reduction of worm burden, in shortening size of adult nematodes and in increasing the number of inhibited larvae [304]. However, the full combination of the three manifestation of immunity are not always simultaneously observed especially because they do not develop at the same rate [304, 489]. In *T. circumcincta* it has been proposed that immunity primarily acts on worm fertility and then on worm burden [489].

In *H. contortus* infection, a "self-cure" phenomenon has been described in which adults are completely swept out the host after massive larvae ingestion [371]. The IgE-mediated hypersensitivity response is thought to be responsible of this reaction that is usually associated to acute diarrhea [371]. The hypersensitivity is also at stake in the rapid rejection of infective larvae [268] and it can also be long lasting hence resulting in granuloma formation subsequently altering the absorption and digestion process.

Even if protection is not complete, immune sheep usually perform better in subsequent infections [461, 507]. After challenge, the CD4+/CD8+ ratio is increased and antibodies are produced in a quicker and stronger fashion [498].

The precise mechanisms leading to the rejection of GIN in sheep are far from understood even if knowledge has been accumulated in recent years. The protective response obtained after

a first challenge could be a potential lead for vaccine development.

1.4 Application to the development of vaccine against sheep nematodes [477]

The ability of sheep to mount an effective acquired response that amplifies as the lamb ages has motivated vaccine development. However until now, no vaccine against ruminant nematodes but DICTOL (that confers protection against *Dictyocaulus viviparus* in cattle) have been released to the market [477, 531, 222]. To be efficient, developed vaccine should at least mimic the basic steps of the immune response that are recognition of parasite antigens, activation of appropriate immune effector mechanisms and expulsion or killing of the parasite [357]. To assess the efficacy of vaccines, both the specific antibody titres and parasitological examination (FEC, worm burden) are determined [498].

So far, the use of natural antigens have been too variable in their on-field efficacy so that research have looked for hidden antigens [462, 533, 212, 210, 387]. To be opposed to the natural antigens, are the hidden antigens that are usually not recognized by the host immune system [280]. Even if these antigens are efficient against blood-feeding parasites [280], the immunity they confer does not hold under pasture conditions thus requiring several injections to promote immunity [308].

As opposed to anthelmintic treatments that produce a complete drench of animals, vaccine would break the epidemiological dynamics of GIN infection by lowering the flock parasitic burden [477]. Modeling *T. colubriformis* populations in grazing sheep, Barnes *et al.* concluded that protecting 80% of the flock would be sufficient to control infection in comparison to a classical anthelmintic program [36].

Commercial interest is considerable in this research area so that much of the data remain unpublished [387]. We propose a review of the different tracks that have been explored so far and the associated results, pitfalls and hurdles that explained the absence of efficient vaccine to date.

1.5 Application to the diagnostic and the monitoring of nematode infection in sheep [130, 498]

The diagnosis of gastroenteritis mediated by GIN result from both clinical and epidemiological considerations. Direct evaluation tools have been developed that can also serve as monitoring parameters for making or not the drenching decision (see next section 2.3.1.2). Ideally, parameters monitoring resistance of sheep should be easy to sample, reliable and repeatable, and their analysis would be able to be automated. Samples should also be storable for laboratory analysis and for repeat assays and preferably unaffected by drenching practice [130].

1.5.1 Clinical approach (table 1.3)

Scoring individuals helps identifying animals requiring anthelmintic treatment through indirect evaluation of parasitic load. However these tools are usually not specific, requires time and additional work load for breeders thus making their implementation difficult [77].

Table 1.3: Indirect indicators of GIN infection

Evaluation method	**Scoring scale**	**Pros**	**Cons**	**Ref**
Weight	variable	Precise measure	Work load , poor specificity	[6]
Body score	0 to 5	Easy to measure	Work load, training required	-
Diarrhoea score (DISCO)	1 to 3	Easy to perform	Seasonal interpretation, prior parasitological knowledge required; To be associated with FEC	[79]
Dag score	0 to 5	Easy to perform	Not specific, visual examination	[295, 296]
FAMACHA	1 to 5	Easy to perform	Training required, work load; For regions where haemonchosis is preponderant	[333]

1.5.2 Laboratory parameters

1.5.2.1 Fecal Egg Count (FEC)

Counting the eggs in a standardized quantity of feces is the most favored laboratory test for GIN infection. This technique is relatively cheap and non-invasive [351]. It is also easy to perform and requires little prior knowledge so that breeders can do it themselves on farm [273]. FEC also gives access to some epidemiological information as species can be differentiated on the basis of egg morphology.

This method consists in diluting 3 g of feces in saturated salted water. GIN eggs hence remain on the liquid surface which facilitates their counting in a McMaster slide. This slide contains two graduated squares in which eggs are counted. Summing the number of eggs counted and after correction for the sampling performed, gives the total number of eggs/g of feces.

However interpretation of FEC is complicated by GIN biology. Indeed the relationship between the total number of worms and FEC is not as straightforward as it was supposed. For instance, *H. contortus* females are high eggs producers whereas *Nematodirus spp.* shed eggs in a discrete manner [560]. A related complication linked to natural challenge is the lack of knowledge about host exposure.

Furthermore FEC also vary with host health, as diahorrea is associated to a dilution of eggs [301] and a poor host body condition can result in lower egg excretion [351].

Still, FEC is a useful tool to evaluate flock contamination in order to better target groups of livestock that require treatment [376]. Composite FEC based on mixed equal amount of feces from several animals can help reducing the associated work load. While doing so, attention should be paid on the varying degree of parasitic aggregation that can lead to underestimation of the parasitic load [376].

1.5.2.2 Immunoglobulin concentration

The determination of Ig concentration in sheep serum could be a good tool to monitor host resistance to parasites. Ig concentration (IgG in *H. contortus* and *T. colubriformis* infections [130], or IgA in *T. circumcincta* infection [109]) has been positively correlated with other immune parameters and negatively associated to FEC.

A recent study validated a salivary anti-CarLA IgA dosing by ELISA to monitor sheep resistance in flocks thus reducing work load in comparison to FEC sampling [468].

Even if of major interest for breeding purpose, Ig dosing does not really provide an information about parasitic load at the time of sampling.

1.5.2.3 Plasmatic pepsinogen concentration

Plasma pepsinogen is a pathophysiological marker of abomasal lesions. Pepsinogen is a precursor of the digestive enzyme pepsin. Any increase in the pH of the abomasum prevents the conversion of pepsinogen into pepsin. It was firstly thought that a combination of pH rise, the accumulation of pepsinogen in gastric glands and an increased permeability were responsible

for pepsinogen concentration increase in blood [164]. However recent findings are not as clear as hyperpepsinogenemia can also be mediated by parasite ES products or host hormone, *e.g.* gastrin.

Nevertheless, pepsinogen concentration can be used as a marker of infection by abomasum resident nematodes like *H. contortus* or *T. circumcincta*, values of infected sheep ranging between 1.5 and 2.5 U Tyr/L [507]. Interestingly, higher basal levels of pepsinogen concentration in blood have been reported in lambs of the resistant Martinik Black-belly breed [507]. This might reflect a higher inflammatory state in this resistant breed and would be in agreement with the view that hyperpesinogenemia results from the host response itself [471].

1.5.2.4 DNA quantitation

Some works have attempted to detect FEC through PCR methods [497]. Sensitivity of the technic was higher than the usual McMaster technique and also provides information about contaminating species but no quantitative information was available on parasitic load [497]. Another work by Bott *et al.* mixed both eggs isolation through salt flotation technique and quantitative PCR targeting ribosomal DNAto provide a semi-quantitative information on parasitic load [63].

1.5.2.5 Worm burden determination

In case of mortality or for research purpose, worm burden can be directly measured. This provides information about species and development stages. A total pathogenic index has been proposed to better assess the pathophysiological consequences of nematode infection [351]. This index weights each count by a factor depending on their pathogenicity or their fecundity, e.g. 500 for *Haemonchus* or 6,000 for *Trichostrongylus spp.*. In the end, indexes are summed. This system has been used by Abbott *et al.* in their SCOPS [2, 501].

Chapter 2

Opportunities for selection of sheep resistant to nematodes

This part discusses the opportunities to use genetics as way of control of gastro-intestinal nematodes in sheep. It is based on the work of Davies *et al.* [107] who assessed the opportunities of using genetic selection for managing nematodes in sheep. In their work, seven different factors were considered ranging from the importance to the sheep breeding industry to socio-economical aspects. Through the reviewing of recent literature we follow the same criteria to justify the breeding of resistant animals.

2.1 GIN are the curse of the (meat) sheep breeding industry

As discussed in the last paragraph, GIN are to be found everywhere sheep are bred, especially in extensive production systems in either temperate or tropical areas. This parasitic load has two major direct consequences: production losses, either through direct mortality or retarded growth, and treatment costs. Production losses have been well described. Some tropical species are particularly pathogenic. For instance an infection dose of 80 larvae of *Oesophagostomum columbianum* can lead to the death of lambs [306]. In addition *H. contortus* infection, also called haemonchosis, can be particularly severe with the development of an anemia and ensuing the death of the animals as shown by Mugambi *et al.* [377]. In temperate regions, local species mostly result in gastro-enteritis in young animals resulting in lower growth rates but also to the death of the lambs through dehydration. One of the most frequently encountered worm, *T. colubriformis*, is known as the "bankrupt worm" which underlies well its economic impact.

In addition the economic impact of uncontrolled parasitism is exemplified the closure of some individual farms in the South Africa or in the UK [55, 524] due to the successive failures of all available broad-spectrum anthelmintics.

Beyond these illustrations, some studies aimed at precisely quantified the costs linked to nematodes for the sheep breeding industry. In 1995, McLeod estimated the total costs of GIN in Australia to be A$222 million [352] with production losses (141 $M) being greater than treatments costs ($ 81M). He also stated *Nematodirus spp.* as the greatest constraint. More recently a report by Sackett *et al.* (2006) ranked GIN as the top economical threat for the sheep industry with a A$ 369 million loss and a cost per head ranging between 2 and 12 A$, hence representing 3 to 20% of the breeder incomes [450]. In Europe, Nieuwhof and Bishop [391] also found GIN to be a great load on the UK sheep industry with an estimated annual cost of £84 million. Both studies also found that reduction in incomes was mostly due to production losses rather than to treatments cost [391, 450]. In New Zealand, Leathwick cited in [317] estimated the cost of anthelmintic resistance in 2002 to be NZ$18 million per annum and predicted an increase up to NZ$ 60 million over the next 20 years.

2.2 Anthelmintics are not a unique sustainable solution any more

Both the ubiquity of sheep nematodes and the production losses they are responsible for is a non negligible threat for sustainable sheep breeding. Efficient ways of control are hence required. Before the advances of modern chemistry, treatment of animals for internal parasites mostly relied on the mechanical removal of the parasites from their predilection sites either with metals or plant extracts [350]. At the beginning of the 20^{th} century some advances were made through the use of crude drugs like arsenicals, nicotine sulfate, copper sulfate, carbon tetrachloride, tetrachloroethylene that were equally toxic for both the parasites and their sheep hosts [350, 456]. The first modern veterinary anthelmintics, *i.e.* drugs acting against helminths, were released after the 1940ś: phenothiazine (1940) and piperazine (1954). These two compounds were subsequently followed by major progress between 1960 and 1980 [350]. This section describes the available molecules for sheep drenching and the associated resistances developed by worms. Available alternatives to anthelmintics are also reviewed.

2.2.1 The state of play : review of available molecules for sheep in France

Anthelmintics are classified into three major classes in function of their chemical structure and properties: benzimidazoles, imidazothiazoles and macrocyclic lactones (see table 2.1). These three classes have a broad spectrum but have different ways of acting on worms [1].

Table 2.1: List of available anthelmintics for sheep in France

Chemical name	**Parasites controlled**	**Administration**	**Target**
		Benzimidazoles	
Albendazole	GIN, lungworm, tapeworm, Fluke	Oral drench	
Fenbendazole	GIN, lungworm, tapeworm	Oral drench	
Mebendazole	GIN, lungworm, tapeworm	Oral drench	
Mebendazole + closantel	GIN, lungworm, tapeworm, fluke, nasal bots	Oral drench	β-tubulin inhibition
Netobimin	GIN, lungworm, tapeworm, fluke	Oral drench	
Oxfendazole	GIN, lungworm, tapeworm	Oral drench	
Oxfendazole + oxyclozanide	GIN, lungworm, tapeworm, fluke	Oral drench	
	Imidazothiazoles \Tetrahydropyrimidines		
Levamisole	GIN, lungworm	Oral drench	
Levamisole	GIN, lungworm	S/c injection	
Levamisole + oxyclozanide	GIN, lungworm, fluke	Oral drench	Nicotinic receptor agonists
Levamisole + triclabendazole	GIN, lungworm, fluke	Oral drench	
Levamisole + praziquantel	GIN, lungworm, tapeworm	Oral drench	
Morantel	GIN	Oral drench	Nicotinic receptor agonists
		Macrocyclic lactones	
Doramectin	GIN, lungworm, ectoparasites	I/m injection	
Ivermectin	GIN, lungworm, nasal bot	Oral drench	
Ivermectin	GIN, lungworm, ectoparasites	S/c injection	Glutamate-gated
Moxidectin	GIN, lungworm	Oral drench	chloride channels
Moxidectin	GIN, lungworm, ectoparasites	S/c injection	
Long Acting Moxidectin	GIN, lungworm, ectoparasites	S/c injection	
		Salicylanilides	
Closantel	Fluke, H.contortus , Oesophagostomum spp, nasal bot	Oral drench	
Closantel + Mebendazole	GIN, lungworm, tapeworm, fluke, nasal bots	Oral drench	Hematophagous parasites
Closantel + Oxfendazole	GIN, lungworm, tapeworm, fluke, nasal bots	Oral drench	
	New molecules (recently or to be released)		
Aminoacétonitrile derivative (Monepantel)	GIN	Oral drench	

After the "electric years" [350], no new anthelmintic class has been introduced to the livestock market within the last 28 years [256]. This issue was reported by Besier as "a hypothetical nematode-control time bomb with a shortening fuse" especially for the sheep and goat industry [43]. One year after, monepantel (AAD1566) was finally released [256], quickly followed by a combination of derquantel and abamectin that was put to the livestock market in 2011 [317].

2.2.2 Resistances of parasites to anthelmintics are frequent

2.2.2.1 Definitions

A definition of anthelmintics resistance is given by Prichard in 1980 : "Resistance is present when there is a greater frequency of individuals within a population able to tolerate doses of compound than in a normal population of the same species and is heritable" [421].

Provided that molecules belonging to the same class have the same way of action, resistance to one of this molecule usually confers resistance to others, also known as "side-resistance" [564]. However such extrapolation are less evident between nematodes species [455].

2.2.2.2 A brief status report (table 2.2)

The first failure of anthelmintic was reported in the late 1950s in *H. contortus* in sheep [260]. Subsequently, resistance appeared almost as quickly as the new drugs had been released, firstly for benzimidazoles in *H. contortus* and then for multiple other species and other drugs [260, 350]. The first reported cases of "Multiple Drug Resistant" (MDR) parasites in the 1980s generalized ten years later and MDR *H. contortus* now threaten the small ruminant industry mostly in tropical and subtropical countries [260, 564, 350, 246]. The extent of the problem is well illustrated by the abundant literature of worldwide case reports (see table 2.2) that are still arising and worsening in the 2000s [261, 7, 473, 457, 532]. The situation is particularly concerning in Australia, New-Zealand and South-Africa as well as in tropical regions like Brazil and Malaysia [246, 260]. On the top of this, this issue is currently extending to cattle and horses [261, 495].

2.2.3 Anthelmintics impact the environment

The last decades have seen a growing concern about ecological impaction of agriculture in the public opinion [349, 324]. It is known that anthelmintics are not completely metabolized within the host and that some active molecules are excreted in the feces and/or urine [349].

Table 2.2: Worldwide reported cases of anthelmintic resistance in sheep (adapted from www.parasol.org)

Country	**BZ**	**LEV**	**ML**	**Reference**
Argentina	+	+	+	[139]
Australia	+	+	+	[403]
Belgium	+	-	-	[530]
Brazil	+	+	+	[138, 7, 84]
Cameroon	+	+	+	[382]
Denmark	+	+	+	[54]
England/Wales	+	+	+	[574]
Ethiopia	+	+	+	[473]
France	+	+	-	[88]
Greece	+	-	-	[406]
India	+	+	-	[183]
Ireland	+	-	-	[394]
Italy	+	-	-	[175]
Kenya	+	+	-	[330]
Malaysia	+	+	+	[85]
Netherlands	+	+	+	[62]
Nigeria	+	+	+	[344]
New Zealand	+	+	+	[302]
Paraguay	+	+	+	[326]
Scotland	+	+	+	[457]
Slovakia	+	-	+	[83]
South Africa	+	+	+	[525]
Spain	+	-	-	[433]
Sweden	+	-	-	[227]
Switzerland	+	-	+	[225, 20]
Tanzania	+	-	-	[53]
Turkey		+		[512]
Uruguay	+	+	+	[381]
USA	+	+	+	[239]
Zambia	+	-	+	[170]
Zimbabwe	+	+	-	[57]

Key: + : resistance has been reported, : resistance has not been reported.

The increased spectrum of activity of macrocyclic lactones, that target both nematodes and arthropods, is synonymous of noticeable insecticidal activity on native pasture fauna [324]. However consensus has still not been reached on this point [324].

2.3 Very few alternatives are efficient and/or available

Following the development of anthelmintic resistance, extensive research has been done on alternatives to the classical drenching regimens. The research in this field is particularly difficult as proposed control options have to be efficient but also easy to implement. They fall into three different categories :

- Drenching in a more efficient way
- Escaping challenge
- Enhancing host response

2.3.1 Drenching in a different way

2.3.1.1 Avoid risk factors while using anthelmintics [470, 564]

Avoiding known risk factors and making a better use of drugs that are available is the first option stakeholders have. It has been suggested to alternate anthelmintic classes on farm [37, 337], to prevent the introduction of resistance by quarantine drenching [564], and to take care to give the right anthelmintic dose. Long acting molecules could also be involved in the apparition of resistances [300, 494, 299].

To do so and to impact farmer practices and attitudes, a working group of the UK, called SCOPS (Sustainable Control Of Parasite in Sheep) was created in 2003. They produced some guidelines intended for vets (available on www.nationalsheep.org.uk) in order to minimize the selection pressure for anthelmintics resistance [501]. So far, such working a group has not been created in France.

2.3.1.2 Refugia-based approach and targeted selective treatment

In the context of veterinary parasitology, refugia is the proportion of nematodes that are not under treatment selective pressure [44, 470, 272]. After these parasites complete their lifecyle

and pass on their genome to the next generation, it is expected that resistant worms will be diluted in a greater proportion of susceptible infective larvae [522] in [44, 272].

This concept has been firstly used by Martin *et al.* [338] and Van Wyk recently recalled this concept as a major tool for the management of anthelmintics resistance [522, 272]. Indeed, it implies that a sub-population of worms should not be exposed to treatment hence being equivalent to applying drenching to targeted animals. This targeted selective treatment approach can lead to a reduction in treatment frequency and/or reducing the number of flocks to be drenched on a same property [44, 272, 273].

Decision criteria can be of different nature, either considering the size of parasite burdens (through fecal egg counts) or evidence of parasitism (thanks to indirect indicators). For this latter case, the FAMACHA© system originally described by Dr. Faffa Malan (**Fa**ffa **Ma**lan **Cha**rt) and Van Wyk in 1992[333] provides a severity grading of haemonchosis based on the inspection of sheep conjunctivae [526]. Animal production parameters, *e.g.* milk yield [102] or bodyweight [303, 481, 168] can also be used [44]. Another and more generalist approach is to consider diarrhea indicators, like the Disco (diarrhea score) and the dag score, that are however less specific [79, 77]. Each parameter has its pro and cons and shows different breeders' acceptance. In addition, choosing appropriate evidence-based threshold for drenching is not obvious: this can be either done by subjective thresholds [172], or by a flock-based threshold [102] associated to the subsequent treatment of outliers [498]. Computer models can also provide additional insights on the way to implement this approach [168, 123]. Other authors suggest drenching at random could also be applied [167].

2.3.1.3 Non conventional treatments

A related alternative to the better use of anthelmintics is the use of non conventional treatments that can be either plant extracts with anthelmintic properties or mineral elements (copper oxide wire particles [281, 72, 71, 70, 73, 478] and multi-trace element slow release device) that can affect the survival/settling of worms [236].

The use of plant extracts primarily comes from the traditional pharmacopea from local ethnoveterinary knowledge that is used in many parts of the world [236, 186], as in Africa [195, 284, 348], in Asia [4, 154] or in Europe [534, 186, 545]. Such knowledge does not always show efficacy and belief may be the only justification to their use [545]. However the screening of their biological activities can also be successful [348] but is usually lacking [348, 154, 284, 195].

In addition to these ethno-veterinary drugs, there is a strong parallel interest in the development of "nutraceuticals" that can be defined as "fodders whose traditional exploitation for animal feed is nowadays associated with a beneficial health effect" [15] in [236]. Bioactive principles responsible for anthelmintic activity seem to be plant secondary metabolites like saponin, alcaloïd, tannins [186] and condensed tannins are reported as having a major role [237]. Precise way of acting of these compounds are still under investigation.

2.3.2 Avoid worm challenge

Other alternatives focus on the reduction of the contact between the host and the parasite's infective stages. This can be either through evasion strategies that aim at reducing the probability that potential hosts meet the pathogens or through the specific hunt of parasitic stages by using predacious fungi.

2.3.2.1 Pasture management

The first approach is to reduce the stocking density on pasture. Indeed, the fewer individuals, the smaller the probability that hosts are going to graze around feces where the contamination is the most likely. This is both the most efficient way of control and certainly not the most economically sustainable approach [487].

A much more elaborated approach had been firstly introduced by Michel [367, 368] that stated three different strategies [236, 35]:

- evasion, relying on the move of animals before critical levels of infection are to be reached
- prevention, relying on exposing animals to safe levels of infective larvae (clean pastures or early anthelmintic treatment
- dilution, aiming at getting the parasite lost into helminthologically inert animals, resistant animals from the same species or different species

Rotational grazing is synonymous of a heavier work load for stakeholders as well as some potential negative consequences to productivity if the grazing system does not aim at taking the best from the fodder resource [236].

2.3.2.2 Removing the worms from pastures

On the opposite of the defensive strategy that represents rotational grazing, are the offensive sanitary measures that consist in directly eliminating worms from pastures. As reported by Hoste and Torres-Acosta (2011), three different types of weapons are available, i.e. chemical, physical or biological [236].

The efficiency of chemical materials (calcic cianamid,urea) was demonstrated under laboratory conditions but did not hold under field conditions. In addition such strategy goes against the actual environmental concerns and does not represent the cheapest way of control [236].

A maybe most important option is the use of biological agents capable of killing the worms like the nematode killing microfungi *Duddingtonia flagrans* [236, 298, 546, 247, 487, 297]. However the lack of consistency in its on farm efficacy in Europe [150, 145] has reduced the interest for this approach [247]. In addition technical issues remain unresolved and thus hamper the complete success of this strategy to jump into the commercial phase [236].

2.3.3 Enhancing the host response

2.3.3.1 Host nutrition

GIN infection is known to have a strong impact on livestock production. This occurs through four different ways that are [282, 97]:

- A reduced voluntary feed intake
- Some endogenous losses, mainly due to the lack of reabsorption, blood losses (haemonchosis) and repair processes
- Reduced absorption of nutrients depending on whether the lesions are in the anterior or the distal tract
- The host immune response that is particularly demanding in proteins, either in mucosal secretions (threonine, serine and proline) or in the production of cytokines that directly competes with wool production for sulfur-amino acids

All these effects vary according to the level of infection but disturbances are usually more severe in the protein metabolism than in the energy balance [97, 282]. These patho-physiological consequences are also aggravated by the high physiological needs of the most heavily infected

hosts, *i.e.* growing lambs and periparturient ewes. This is the support of the general "partitioning framework" developped by Coop and Kyriazakis that orders nutritional priorities in a parasitised host [96]. This framework has been reproduced in table 2.3.

Table 2.3: Partionning framework of host nutrients for various physiological states (reproduced from [96])

Growing animal		Reproducing animal
Acquisition phase	Expression phase	
1. Maintenance of body protein	1. Maintenance of body protein	1. Maintenance of body protein
2. Acquisition of immunity	2. Protein gain	2. Reproductive effort
3. Protein gain	3. Expression of immunity	3. Expression of immunity
4. Maintenance and gain of body lipid	4. Maintenance and gain of body lipid	4. Attainment of desired fatness

The top priority is the body maintenance that conditions the animal short-term survival, immediately followed by the ensuring of the preservation of the animal's genetic material [96]. It follows that optimizing host nutrition may increase the ability of the host to cope with parasites or to contain it [96, 97]. This theoretical framework has been supported by various studies reviewed by Knox *et al.* (2006) [282].

The point to be adressed again is the practical implementation of such strategy. The required supplementation and the frequency of the treatment have to meet a balance between efficiency and the associated costs and labour it generates [282].

2.3.3.2 Vaccination ?

This topic has already been discussed in a previous dedicated part (1.4).

2.3.3.3 Selecting animals that better resist the infection

Modern breeding techniques have concentrated animals on restricted areas and have focused on selecting highest producers to increase incomes. Combined together these two factors have increased the infection pressure and have selected for animals that require more inputs and high-level management practices [416, 112]. However some variation exists within flock and parasites are over-dispersed: some rare individuals concentrating most of the parasitic load. Therefore it should be possible to select for more resistant animals that could better endure parasites.

This subsection aims at discussing the concepts underlying the *"resistance"* term. This notion can be split into several related aspects, that are resistance, tolerance and resilience [49] and hosts have evolved both resistance and tolerance [429]. Resistance is defined as the ability

to limit parasite burdens, either through reduction of the development rate of the parasite or its clearance or both [120]. On the contrary, tolerance is the ability to cope with the pathogen by limiting its detrimental effects [429, 120]. Resilience is a related aspect of tolerance that also considers the production level of the animal under challenge [49]: a resilient animal is able to maintain its production level while being infected [6, 51].

This dichotomy not only helps understanding underlying mechanisms of the host response but also provides insights on evolutionary aspects as tolerance put less or no pressure on the parasite [429]. Therefore resistance is thought to select for parasites able to overcome host defence leading to antagonistic co-evolution between host and parasite [429, 568]. This aspect has been also known as the "Red Queen theory" [521]. The theory is based on the observation to Alice by the Red Queen in Lewis Carroll's "Through the Looking Glass" that "*in this place it takes all the running you can do, to keep in the same place*", which is similar to the amount of efforts required for both host and pathogens for maintenance [521, 320]. Experimental evidences of this theory have been lacking [117] but some recent works have brought insights to it [117, 465]. Decaestecker *et al.* mined pond sediment and focused on an aquatic flea and its microparasites to demonstrate that the parasite rapidly adapted to its host (within a few years), simultaneously loosing infectivity against past genotypes [117]. In another study using the free living nematode *C. elegans* and a bacteria, *Bacillus thuringiensis*, Schulter *et al.* demonstrated an increase in genetic diversity over time associated to adaptation costs in both parasites and host after 48 generations [465].

While resistance is typically measured as the inverse of the infection density (no. parasite per host) [429], resilience and tolerance are somewhat harder to assess as they correspond to a rate of change in fitness over parasite infection [51, 429]. Bisset and Morris distinguish three measures that can be the growth rate depression under infection, the growth rate in comparison to peers' performance or the number of drenches required by an individual [51]. Raberg reviewed the two concepts in a single equation :

$$W_i = a_i + b_i I$$

where W_i is fitness of hosts of type i, a_i is the intercept (*i.e.* the fitness when uninfected), I is infection intensity, and b_i is the slope of the relationship between $W$and I, that is, tolerance.

Variations in I is synonymous of resistance whereas variation in tolerance will be assessed by

changes in the slope b_i [429]. The practical consequences of this difference is that the measure of tolerance cannot be done on a single individual but on different genetic groups challenged with various parasite loads hence giving access to the slope value [429].

In summary, defining the notion of resistance is not only a semantic matter as it is associated to various outcomes, especially in the frame of genetic selection of resistant animals (developed below in section 3.2).

2.4 Summary on the available alternatives

In conclusion, there has been a strong research effort towards the development of efficient alternative ways to control GIN infection in sheep in the last few years. Among these, rare projects have been successful alone, either because of the intrinsic efficiency of the approach or because of the work load it is synonymous of. This hence underlies the necessity of an integrative management approach such as what has been done in plants.

In addition to all of these alternatives, genetics may have a key role. Breeding resistant animals would have a double impact with both the enhancement of the host resistance as well as a reduction of the environment contamination (as the more resistant are the animals, the less eggs they will shed on pasture).

Some illustrations of this management strategy and associated options will be developed in the two following chapters. A first chapter will be dedicated to the classical genetic selection approaches whereas the use of molecular information to improve sheep resistance will be developed in two different chapters, one focusing on candidate gene approaches and the other on the identification of genetic markers explaining significant part of the total genetic variance.

Chapter 3

Classical selection of resistant animals

The use of genetic strategies for improving health status of sheep flock still faces some arguments that deny the efficacy of such an approach. Indeed, quantitative genetics theory predicts that any gene with a favorable effect on host survival will be passed through generations and should finally be fixed by natural selection [485]. Therefore no genetic variation should be observed for traits influencing resistance to diseases. This chapter aims at showing that genetic variation still exists for resistance to nematode infection and that it can provide an efficient alternative way of managing nematodes infection. Some examples of breed comparison are given as well as evidence of genetic variation within breeds of sheep.

3.1 Using the breeds that are already adapted

One of the simplest use of genetics as an alternative way of control of GIN infection is to choose the appropriate breed for the production environment.

3.1.1 Some examples of breed susceptibility differences

There has been a great effort in documenting breed differences in resistance to nematode infections (reviewed in [112, 416, 49, 485]) and some results are given in table 3.1. Great differences have been reported across production systems, climatic areas and for a wide range of parasites [112].

A consensus has emerged on the better adaptation of "tropical" breeds over the european

Table 3.1: Between breeds comparison for resistance to GIN infection

Resistant breed	Compared breed	Location	Reference
Gulf Coast Native	Suffolk	USA	[369, 370, 315]
St-Croix x Sumatra	Sumatra	Indonesia	[441]
Java Fat-tail x Sumatra	Sumatra	Indonesia	[441]
Barbados Blackbelly x Sumatra	Sumatra	Indonesia	[441]
Menz	Horro	Ethiopia	[503, 431, 206]
Red Maasai	Dorper	Kenya	[378, 549, 28]
Red Maasai	Dorper, Romney, Blackheaded Somali	Kenya	[377]
Lohi	Kachhi, Thalli	Pakistan	[451]
Criollo	Suffolk	Mexico	[5]
Santa Ines (pure breed, cross bred)	Suffolk, Ile de France	Brazil	[12, 13]
Local Kashmiri	Crossbred Kashmir Merino, Bhakarwal, Corriedale	India	[500]
St-Croix hair	Wool (50% Dorset, 25% Rambouillet, 25% Finnsheep)	USA	[327]
Canaria Hair	Canaria	Spain	[192]
Garole sheep	Decanni, Bannur	India	[392]
Blackbelly	Romane	France	[24, 198]
Djallonke	Djallonge-Sahelian	Gambia	[194]
Merinoland	Rhoen	Germany	[173, 229]
Dorper, Katahdin, St-Croix	Hamphsire, Suffolk	USA	[69, 68]
Florida Native	Rambouillet	USA	[10]
Sabi	Dorper	Zimbabwe	[341]
Texel	Suffolk	UK	[193]

breeds [49, 112]. For instance, breeds from the Caribbean or from Africa show better resistance and better ability for production in area with high parasitic load (see table 3.1). This has been well illustrated by Mugambi *et al.* showed the extreme susceptibility of the imported Romney breed when put under field infection in Kenya [377]. After a first natural infection running on one year, nine out of the 15 Romney infected males died from haemonchosis whereas other local breeds (Dorper and Blackheaded Somali) were less impacted and that every Red Maasai stayed alive. Red Maasai superiority over Dorper and Blackheaded Somali during haemonchosis was subsequently confirmed in an artificial challenge [377]. A subsequent comprehensive study with 212 Red Maasai and 311 Dorper lambs showed the Red Maasai breed had significant lower death rate from birth to one year of age [28]. Other experiments by Burke and Miller [69], Amarante *et al.* [12, 11, 13] or Aumont *et al.* [24] have also demonstrated the better ability of local Caribbean breeds, *i.e.* St. Croix, Santa Ines or Martinik Blackbelly respectively, to cope with parasitic load. Interestingly the superiority of the Barbados Blackbelly and St. Croix was still observed and even better than the intermediate status of the Indonesian Thin Tail breed when put under indonesian condidtions [492] in [49].

It could be argued that productive breeds from European countries are not adapted to tropical conditions, and that the observed differences only reflect the change in the environment. However resistance differences have not only been observed between breeds from different climatic areas. For instance, a 3-year study by Good *et al.* performed in the UK showed that Texel lambs and ewes carried less worms and excreted less eggs than than the Suffolk individuals [193]. In addition Gonzalez *et al.* also demonstrated some differences in FEC between two local breeds of the Canary Islands, the Canaria Hair Breed used for milk production being less resistant than the Canaria sheep dedicated to meat production [192] and three Pakistani sheep breeds also showed different susceptibilities to *H. contortus* [451].

3.1.2 Using the heterosis effect

Even if resistant breeds are perfectly adapted to heavy parasitic load, their production performances are poor. For instance, 1-year old Suffolk lambs weighed around 20 Kg more than their Native counterparts under USA conditions [315]. In addition to their small format, resistant breeds are usually hair breeds which is not consistent with wool production (table 3.1).

Therefore the use of adapted breeds is not an optimal solution as breeders would only go from dying sheep to healthy non producing animals which are, in terms of economic survival,

quite synonymous situations. A possible option is to take advantage from both resistant and high-production breeds by cross-breeding. Such cross takes advantage of what is known as the "heterosis" effect, that is the average superiority of the offspring in comparison to the mean performance of the parental breeds [151]. This "hybrid vigor" is usually greater for traits related to fitness [151] which could be of interest in the frame of breeding animals resistance to GIN infection. While comparing different breeds performances under infection, a few studies focused on this heterosis effect [9, 10, 198, 315, 28, 229].

In a 6-year long study, Baker *et al.* surveyed performances of 1785 lambs with different genetic background, *i.e.* straight Dorper and Red Maasai, cross bred and back-crossed lines form a diallel design. Animals were surveyed over one year and regurlarly sampled for PCV and FEC determination and weighing. Following the approach described by Robison [439], the genetic basis of crossbreeding was split into additive and heterosis genetic effects from the individual and its parents, that were applied weighing coefficients indicating the contribution of each breed to the genetic background of the considered individual (i.e. progeny, sire or dam). In the end, none of the heterosis estimates was found significant for any of the measured traits [28].

In another study, Li *et al.* used a more simple approach and estimated the mean heterosis effect between Suffolk and Native breeds (as $H = \frac{F1 - \frac{Suffolk+Native}{2}}{\frac{Suffolk+Native}{2}} \times 100$) in a 253 lambs population with half pure breed lambs and half cross bred lambs [315]. They found heterosis coefficients ranging from 33.8 to 81.6% for FEC (according to the age of lambs) and somewhat lower for PCV (0 to 20.7%). These estimates contrast with the non significant effects of Baker *et al.* as well as those of Hielscher *et al.* who estimated heterosis values ranging between -2.1 to 0.5 % for FEC with Rhoen and Merino Land sheep populations [229]. In all cases the crossing of animals produced heavier crossbred lambs with heterosis effect ranging between 1 and 4% [315, 229].

From these limited number of studies it appears that crossing breeds with opposite production and resistance phenotypes could help maintaining sheep breeding under heavy parasite burden whilst getting incomes through intermediary level of production. However such breeding scheme requires to maintain two different selection nucleus which multiplies costs and labor. In addition, it seems to be difficult to implement for wool production purpose as most of the resistant breeds are usually colored-hair breeds [262].

3.2 Using the within breed genetic variation

3.2.1 Genetic variability exists for resistance to GIN

Selecting animals is synonymous to reducing the genetic variation to shift the population level towards better production level. Therefore the available genetic variability needs to be estimated before launching any selection scheme as it conditions the putative genetic gain [151]. This variability is estimated by the heritability coefficient that represents the proportion of the observed variation that is due to genetic variation [151]:

$$h^2 = \frac{var(A)}{var(P)}$$

, where A h^2 is the heritability, var(A), the additive genetic variance and var(P), the observed variation.

Early works by Gregory [197] and Emik [144] (cited in [112]) have demonstrated genetic variation for resistance to GIN in sheep. Since 1980's an abundant literature has been produced on heritability estimates for GIN resistance in broad range of environments and sheep populations [112]. These estimates usually range between 0.2 and 0.4 hence suggesting that genetic progress is possible [49].

Heritability estimates tend to increase as the lambs get older [566, 46, 45, 419]. This correlates with the dam genetic effect that decreases as the lamb get older to become null after 3 months of age and might reflect that little of the genetic variation is due to the innate immune response [566, 46, 45, 563].

3.2.2 How to implement resistance to GIN in selection schemes ?

The first step in setting up a breeding scheme is to define the objective to be achieved and the optimal phenotypic marker for the trait of interest.

3.2.2.1 What goal ?

In the particular case of selecting animals able to cope with GIN infection, the selection objective can be either selecting animals maintaining their production level while infected, selecting animals that get rid of parasites and maintain their production level or simply the reduction of treatments applied [565]. Resistance, resilience and tolerance have been discussed above (see

2.3.3.3). However each strategy has its own costs and returns as discussed by Woolaston and Baker [565].

The main features of selecting for resistance are direct increases of breeders' incomes through production gain and reduced chemical need, as well as the reduction of the pasture infection [46, 565]. A potential drawback is the selection pressure put on the parasite that could lead to the reduction of the refugia and might select for parasites able to withstand host resistance [565, 485]. This point is still controversial. Windon reported nematodes adapted to sheep selected for low FEC [558], but two subsequent studies by Woolaston *et al.* [567] and Kemper *et al.* [269] demonstrated no adaptation of the parasite to selected sheep.

Breeding for resilience could be an alternative as limited selection pressure would be put on parasite [51, 485, 565]. It seems however that disadvantages from such an approach usually outweighs its benefits [565]. Firstly, it appears to be difficult to precisely define objective parameters to be recorded for measuring resilience. Bisset and Morris distinguish three measures that can be the growth rate depression under infection, the growth rate in comparison to peers performances or the number of drenches required by an individual [51]. The growth rate depression does not confound the genetic for growth *per se* as the comparative growth rate does. However this requires two growth rates determination, one while infected and the other while not. Putting the selection effort on the growth rate under challenging environment is what has been usually done before any efficient anthelmintic treatment were available [51]. In addition, a simulation study by Van der Waaij *et al.* showed that under a constant infection pressure, selecting for the observed production would result in simultaneously increasing the resistance level so that it should not be necessary to measure the level of resistance in order to increase it [519]. Secondly, it seems that heritability estimates for resilience were lower than the usual 0.3 value. An early attempt by Albers *et al.* aimed at investigating the genetic relationship between resistance, resilience and production traits [6]. They assessed resilience by the both growth and wool decreases. In the end, heritability estimates were not significantly different from 0 while resilient sires tended to be resistant as well even if correlations had high standard errors [6]. Ten years later, the same pattern was reported by Bisset and Morris [51] who investigated the feasibility of breeding for resilience in a 14,000-lamb population. In that case, the resilience was assessed by the number of drenches required in an individual "on demand" treatment procedure, being either the age at first treatment, the number of drenches or a binary score translating the need or not of a treatment at each drenching point [51]. They showed

that progeny-testing for resilience could increase lambs growth rates and lower dag-score and found significant low heritability estimates, ranging between 0.1 to 0.19 [51]. These low values highlight the difficulty to objectively assess the need of drenching treatment [565]. A positive side effect is that breeders will have to frequently inspect their stock and will develop a greater awareness of when treatment is required thus reducing the undue use of anthelmintics [565]. However it seems rather difficult to implement this scheme under extensive grazing conditions. Thirdly, breeding for resilience does not provide any epidemiological benefit [565].

3.2.2.2 How to evaluate candidates ?

Also it is necessary that selection candidates and their progenies are recorded for resistance to GIN. The major advantage of a natural infection is the huge number of available data. However the impossibility to precisely characterize the infection that occurred (unknown time and dose), and high dependence of parasite cycle to weather conditions can be problematic. Experimental infection overcomes these problems as it is possible to know and to control precisely experimental parameters. In addition, in controlled and well-equipped facilities, it is possible to measure more complex traits than on-field conditions. Nevertheless some questions still remain. Firstly it is costly in terms of time and work load. Secondly, the measured trait may not translate what occurs in natural conditions and it needs to be compared to other results. A study by Gruner *et al.* characterized the response of sheep selected for response to artificial challenge under natural infection conditions [201]. Two different Romanov flocks were artificially challenged with *T.circumcincta* or left on pasture. Extreme sires were subsequently selected and their progenies were either artificially or naturally challenged. In the end, the authors reported high genetic correlation of 0.87 between FEC measured under artificial or natural challenges [201]. These results were in good agreement with other published estimates by Woolaston and Eady who reported a genetic correlation of 0.72 between indoor pen-tested and paddock-tested animals [415] as well as other results reviewed by Woolaston [565]. This is of particular interest as it shows the feasibility of using standardized and simple artificial infection procedures to evaluate genetic merit of rams to withstand a natural challenge.

However the question of which species to employ to evaluate candidates remains as the candidates to selection or their relatives need to be challenged by the appropriate way in a standardized way so that breeding values could be compared [565]. Evidences of cross resistances have been accumulated [565]. Aumont *et al.* demonstrated that genetic correlations between

infections by *T. colubriformis* and *H.contortus* had a genetic correlation of almost 1 [199], hence confirming the observations by Gruner *et al.* [200]. Some more recent studies modulated these findings especially in the case of infection by *T. circumcincta* [263, 271]. Kemper *et al.* found a 93% reduction of *T. colubriformis* burden but only a reduction of *T. circumcincta* development in resistant sheep line selected for 25 years (Rylington Merino flock [262]) in comparison to control [271]. Genetic correlations between Strongyles and *Nematodirus* are moderate to strong according to the season, suggesting some little differences between the two phenomena [46, 563]. The observed differences in that latter case might also reflect the differences in worm biology as *Nematodirus* is known to have a very low fertility [560, 351].

Last but not least is the phenotypic marker to be recorded. Recording resistance is based on the monitoring of pertinent patho-physiological parameters. The best criteria must be discriminant between healthy and infected animals and must show genetic variation in order to be employed in selection schemes. Ideally, this trait has to be easy to measure and cheap [446]. FEC has been the consensual parameter to record individuals susceptibility while a broad range of additional indirect parameters have been developed (described in 1.5). It is relatively easy to perform apart from the work load the egg counting represents and of moderate cost. However one should pay particular attention while interpreting results, as biology can vary between pathogen species (*Nematodirus spp.* and other Strongyles [351]. Clinical symptoms and necropsy lesions can be considered to monitor resistance, necropsy being only dedicated to research purpose. However, these parameters are not typical of only one disease or pathogen and are strongly linked to the clinician's skills, as exemplified by works performed with the FAMACHA© system [526, 67]. Still, genetic parameters estimates for FAMACHA© score are in the usual range of values of resistance traits, *i.e.* between 0.14 and 0.21 [435]. But genetic correlations with classical FEC and hematocrit traits monitored under *Haemonchus* challenge were low and varied according to the season of FAMACHA scoring [435].

Very recently, a novel method has been proposed by Shaw *et al.* that measures the salivary anti-CarLA IgA antibody by swabbing the cheek pouch [468].It offers the advantages to be easy to perform, to avoid withdrawing anthelmintic treatment to know resistance status and to sample every animals at once which is not always the case with fecal sampling [468]. The IgA had better genetic correlations with FEC than IgG and heritability estimates were similar to those for IgG concentration but standard errors were rather high [468]. Even if more practical than FEC, IgA measure may not really fit to the resistance status as it is difficult to differentiate high titers

between basal high IgA producers and high IgA responders. A possible way to increase the accuracy of breeding values is to perform repeated measures as the heritability of n measures is [151]:

$$h_n^2 = h^2 \times (\frac{n}{1 + (n-1) \times r})$$

where r is the repeatability.

From this it follows that the lower the repeatability, the higher the gain obtained through repeated measures. In their study, Gruner *et al.* reported repeatability values ranging from 0.22 to 0.51 between sampling periods and being somewhat higher within periods (from 0.49 to 0.70) [201]. This is in good agreement with other findings [565] and also fits well the reality as worm burdens taken several days apart are more similar than from year to year [201, 565]. Interestingly, Gruner *et al.*'s findings also highlighted a higher repeatability of FEC under natural challenge in comparison to FEC after experimental infection.

3.2.3 What is to be expected from this genetic improvement ?

3.2.3.1 Empirical studies

A recent review of genetic parameters estimates in sheep by Safari *et al.* reported that genetic correlations between worm resistance and production traits were generally negative, *i.e.* favorable, except for birth weight for which the authors reported only one estimate [452]. Another trend proposed by Stear *et al.* was that these correlations tend to be unfavorable in the southern hemisphere (with liveweight and wool production) but favorable in Europe between FEC and liveweight [485]. For instance particularly high favorable genetic correlations of -0.8 and -0.6 between FEC and liveweight were reported by Bishop *et al.* in a Scottish Blackface population [45] and Bouix *et al.* in a Polish breed [64] respectively. This pattern was confirmed in subsequent reports but with estimates closer to neutral relationship, either favorable [46] or null [419, 563].

All these data were obtained in meat and wool breeds and very little is known about genetic correlation between dairy traits and resistance to parasites [112, 81]. In a recent paper, Gutiérrez-gil *et al.* estimated genetic parameters for resistance to GIN in a commercial Churra population but correlations with dairy performances ranged from -0.08 and -0.18 and were not significant [204] thus corroborating findings of others [81].

Most of the knowledge about selection response for resistance to GIN infection comes from

Australia and New-Zealand as reviewed by Karlsson and Greeff [263]. Experimental flocks have been created in both countries like the Rylington Merino flock that has been selected for 25 years. In addition breeding values for resistance to GIN are now available in Australia through *"Sheep Genetics"* (www.sheepgenetics.org.au) and in New-Zealand with the *"Sheep Improvement Limited"* (www.sil.co.nz). Under these southern hemisphere conditions, no negative associated response in production traits could be observed but a higher incidence of diarrhea in the Rylington Merino flock that has been selected for about 25 years [263].

Additional works investigating the impact of this selection on production have been performed [318, 132, 131]. In a first study, Liu *et al.* compared two lines of Merino, selected for GIN resistance and control [318]. They found no difference in feed intake, digestibility and efficiency of utilization of dry matter and nitrogen between the two groups and subsequently no difference in body weight. Under parasite challenge, resistant lambs responded rapidly and strongly with a declined wool production rate but a faster recovery which in the end makes no difference in the overall wool production between the two lines [318]. Doyle *et al.* who compared two divergent lines selected for increased or decreased resistance to *H.contortus* infection, confirmed that both resistance did not affect voluntary feed intake neither diet selection nor production traits (wool and growth) [132]. In another experiment monitoring the rumen physiology, the authors showed that the resistant line exhibited an altered rumen function with increased rate of fluid outflow that might reflect neuromucosal changes linked to the immune response [131].

Other investigations aimed at observing environmental consequences of grazing resistant animals. Bisset *et al.* followed two groups of resistant or susceptible flocks either grazed together or as separate flocks [52]. Resistant lambs grazed together were clearly advantaged over their susceptible counterparts (5 to 6 fold greater contamination pasture and more than 50-fold greater FEC in susceptible line). The same pattern was also observed in Romane lambs grazed on separate pastures during four years [200]. Differences in FEC were still observed under mixed grazing but dag scores were higher in resistant animals, which is not specific of parasitism, and no difference could be observed between the two lines thus making difficult to ascertain the benefit of the genetic selection performed [52].

3.2.3.2 Modeling the outcomes of selecting more resistant animals

To the previous observations mentioned can be added results from *in silico* studies. Modeling host-pathogen interactions can help confirming results of experiment that are often limited by physical boundaries and can also lead the way for conceiving new hypotheses to be tested in experiments [125]. For instance, early work by Barger [33] predicted that selecting for resistance to GIN should reduce pasture contamination and then the frequency of anthelmintic treatments. Progress made through genetic selection was predicted to be greater if epidemiology was incorporated into the model [482]. In addition, Bishop and Stear [47] confirmed the previously described favorable genetic correlation between performances and resistance. Works by Vagenas [518] and Doeschl-Wilson [126] investigated the influence of diet, genotype for growth and resistance and their genetic relationships on genetic parameters estimations. Their models showed that the variability observed between genetic parameters estimates is more likely to be due to the listed environmental effects and their genetic correlations than to host genetic factors.

Gicheha *et al.* published an economic assessment of the efficiency of alternative breeding schemes including resistance to GIN in the index of meat sheep in Kenya [181]. Two different breeding objectives were defined to fit the smallholder and pastoral systems in Kenya. In the first case, the size of the flock or feeding resources were limiting, while in the other traditionnal system sheep perform both tangible (meat, convertible capital) and non tangible goals (socio-economical roles like gifting or insurance) [181]. Whatever the system was, including resistance to GIN into the breeding objectives resulted in increasing the profit per ewe. The authors predicted that sampling 5 to 10 % of the rams according to the system could increase these benefits.

Altogether these findings suggest that selection for more resistant animals should bring benefits only. In addition, Eady *et al.* compared the impact of genetic and non-genetic strategies (experimental vaccination, protein supplementation and strategic drenching) on the control of GIN infection and concluded that the largest and most persistent effect on FEC (69% reduction with genetic selection, *i.e.* two fold the reduction obtained with other strategies) and pasture contamination was achieved through genetic selection [137].

As a conclusion, selection for sheep being more resistant to GIN has been demonstrated to be an efficient solution. Questions remain about its sustainability even if worms do not seem to adapt against resistant selected sheep in experimental conditions. In addition to this sustainability issue is the practical implementation of this breeding objective in selection index

and the associated benefits.

Chapter 4

Is there any Gene Assisted Selection opportunity ?

4.1 What genes ?

Knowledge about phenotypes of interest can help targeting some "functional candidate genes", *i.e.* genes affecting the trait. In the case of resistance to disease, it seems straightforward to have a look at segregation of genes controlling the immune response. These genes can be separated in three different categories: genes controlling innate immunity, others governing the specificity of the immune response and finally genes affecting the quality of the response [25]. This candidate gene approach has been used for resistance to nematodes in sheep, the interferon-γ (denoted *IFN*γ) and the Major Histocompatibility Complex (denoted *MHC*) genes being plausible candidates [486].

4.2 The *MHC* locus

Specificity of the acquired immunity is known to be under control of the MHC complex (see section 1.3), and many studies in pig, chicken, cattle and horse have shown associations of disease resistance with it [25]. However, in comparison to the cited species *MHC* in sheep is poorly characterized [135]. Three classes I,II and III have been characterized, the class II being the best known [135]. Proteins coded by *MHC* class I and II play a major role in antigen processing. In addition, OAR20 which contains *ovar-MHC* locus has been regularly detected as a major region in resistance to nematodes (see following section 5.5).

There is an abundant literature about association of *MHC* locus markers and GIN infection parameters in sheep as reviewed by Hohenhaus & Outteridge [235] and Lee *et al.* [304].

Early works were attempted at CSIRO laboratories in Australia by Outteridge *et al.* [235]. Using sheep lines selected for high or low response to vaccination to *T. colubriformis*, they firstly demonstrated that a particular serotype of an Ovine Lymphocyte Antigen, namely SY1, was differential between the high and low responders and could be split into two subclasses SY 1a and 1b [400]. Subsequently, they found a favorable effect of these two types whereas the SY2 type was associated to higher FEC [402]. To get rid of the population structure associated to the working sheep lines, they performed an analysis in an outbred sheep population, disconnected from the initially selected lines. Their investigation showed that SY 1 groups FEC were halved in comparison to the five groups of lambs of other OLA types, hence confirming the association between the SY OLA type and *T.colubriformis* infection [401]. This success story was more difficult to prove for resistance to *H. contortus* [235]. A first attempt by Cooper *et al.* did not produce significant association between OLA antigen and FEC [98] whereas Luffau *et al.* found a putative effect of the MHC region on FEC in a Romanov flock [323].

For the MHC class II, serological studies could not be performed as in the case of MHC class I and genetic markers have thus been developed by cross-hybridizing of human MHC class II cDNA probes to sheep DNA [235, 242]. One of the first study failed at finding significant association of RFLP markers of *DRB* (class IIa), *DQA* or *DQB* loci (class IIa) with GIN infection [242]. In two complementary studies Schwaiger *et al.* [467] and Buitkamp *et al.* [76] looked for associations between infection parameters and microsatellites developed within the *MHC* locus. Schwaiger *et al.* screened more than 200 Scottish Black Face lambs for polymorphism within the second exon of the *MHC-DRB1* locus. Among the 19 encountered alleles, the authors found that the substitution of the G2 allele to the I allele significantly reduced FEC (by a 58 fold factor) measured in 6 months old lambs under natural challenge (with predominant *T. circumcincta*) [467]. They estimated that this locus explained from 5 to 10% of the observed variation according to the considered time frame (september or october respectively) [467]. Subsequently, Buitkamp *et al.* focused on one class I microsatellite and the class IIb *DY* locus [76]. Even if both loci showed significant association, *DY* showed a considerable effect on FEC reduction but the authors could not conclude if the causative gene lay in between them or if one of them was the searched gene [76].

Stear et al. [488] findings were in favour of an association between resistance to *T.circumcincta*

and *DRB1*, with homozygous sheep having higher FEC than heterozygous. This pattern between homo- and heterozygous individuals was also consistent with Paterson's conclusions [408], whose field study in unmanaged Soay sheep indicated that particular MHC molecules confer either increased or decreased juvenile resistance against strongyles. Using markers within the *DRB* locus, they also evidenced balancing selection, *i.e.* maintenance of allelic diversity, with an excess of nonsynonymous versus synonymous mutation at this locus [407]. This finding was confirmed later on by Charbonnel *et al.* [86] but no QTL could be found on OAR20 in this population [42]. Additional associations between *DRB1*, *DRB2* and *MHC* class I microsatellites and infection parameters of haemonchosis (FEC, PCV, eosinophils concentration) were reported in a Pelibuey sheep population but in that case significant associations usually concerned few individuals (from 3 to 10 carriers) thus making it difficult to interpret the results [82]. Still focusing on the *MHC* class IIa, Hassan *et al.* studied patho-physiological parameters associated to the G2 allele, now denoted *DBR1*1101* allele, carriers at the *DRB1* locus during infection by *T. circumcincta* [215, 214]. The authors used 18 twin pairs, one lamb carrying the *DBR1*1101* allele while the other did not. Lambs underwent *T. circumincta* infection and were slaughtered on different dates corresponding to parasite development stages (days 0, 3, 7, 21, and 35 of the experiment) for parasite examination and tissue sampling for gene expression study. *DBR1*1101* carriers showed lower worm burden, higher mast cell and platelet counts as well as a slow increase in antibody concentration (IgA and IgE) but no difference in FEC could be observed [214]. This acquired resistant phenotype thus seemed more relying on worm expulsion than female fertility. Measures of cytokine gene expression showed a different time course between genotypes [215]. The *DRB1*1101* carriers showed early up-regulation of Th1-associated cytokine quickly replaced by a Th-2 biased response with up-regulation of T-reg cytokines. These changes were observed in non carriers but with a 4-day delay that may explain the observed phenotypic differences between both genotypes [215].

Additional works have focused on the associated *DQ* locus of the MHC class IIa region [266, 265, 228, 163]. Subsequently to the development of an ovine microarray [267] (see section 4.5), Keane *et al.* identified the *DQA1* allele being over-expressed in genetically resistant Perendale sheep while the DQA1 null allele was more frequent in the susceptible line [266]. However the association between this DQA1 null allele and susceptibility to GIN infection was only found in one out of the three breeds tested for validation [266]. This lack of universality was also demonstrated in another experiment that reported an opposite trend as the only significant

association of the DQA1 null allele was with reduced FEC in the South African Meat Merino. This might be due to the DQA1 allele being in linkage disequilibrium with the causative region, the linkage phase differing between breeds [266].

In conclusion, many research works have been performed to decipher the role of *MHC* locus in resistance to GIN. Undoubtedly, this region is pivotal in resistance to GIN. However, it is still difficult to propose a clear way of action as well as a clear role for each locus within this region as reported associations seem to vary depending on the considered breed [458, 40].

4.3 The *IFN*γ locus

OAR3 is one of the most encountered QTL when studying resistance to nematode, and detected regions include (amongst many others) the *IFN*γ locus [127, 49] (see next section 5.5 on QTL detection).

In the same free living Soay sheep population as Paterson [408], Coltman *et al.* considered the segregation of the $IFN\gamma$ to assess the putative role of this region on *T. circumcincta* infection and associated survival rate of sheep [94]. They genotyped two microsatellites flanking the locus and one marker within the first intronic of the gene in both 4-month old and 16-month old lambs for which they measured FEC and IgA concentration. Results showed that markers were in strong linkage disequilibrium but only the intronic marker was associated to FEC and IgA. Association for FEC held for both considered age but the marker was only associated to IgA in young lambs. In addition no association was found with lambs survival [94]. Without providing firm conclusions, authors hypothesized that associated reduction in FEC might be due to a reduction of worm fertility that do not provide survival advantage or that the beneficial effect of the *IFN*γ locus might result in an unbalanced Th2 response that hampers development of efficient Th1 response [94].

Subsequently, Sayers et al. [459] enlightened an association between the *IFN*γ locus and FEC trait for the Texel breed but not in the Suffolk breed. The genotyping of one additional SNP within the *IFN*γ locus confirmed the results. It thus seems that, as what has been observed for the *MHC* locus, genetic association varies according to the considered breed. The authors confirmed this point by studying the impact of the *DBR1* locus in both breeds and found an opposite pattern, with an association found in the Suffolk but not in the Texel breed [458].

Additional work by Dervishi *et al.* studied the *IFN*γ locus structure and developed new

markers within the locus to test for association with infection parameters in naturally challenged ewes with extreme performances (10% highest and lowest FEC) [119]. One microsatellite and twelve SNPs were found (3 SNPs in the coding region, 2 SNPs in the 5' UTR, 5 SNPs in the promoter region and 3 intronic polymorphisms in intron 1 and 1 SNP in intron 3). None of the markers taken alone was associated to FEC but the haplotype considering one SNP and the microsatellite marker was significantly associated to worm burden and FEC. However no difference in *IFN*γ expression could be found between both groups of ewes. This finding associated to the presence of two homozygotes animals carrying the susceptible haplotype allele showing null FEC, led the authors to conclude that a QTL was linked to the *IFN*γ locus. However resistance to GIN infection is known to be under the control of many genes (see sections 3.2 and 5.6.2) that were not considered in their analyses.

4.4 Cytokines and Ig coding genes

IgE is known to have a major role in immune response against parasites but Clarke et al. [91] could not find an unequivocal association between IgE polymorphisms and nematode resistance. They reported a significant association between an IgE allele with *T. colubriformis* infection but could not validate this association in another flock neither for *T. colubriformis* nor for *H. contortus*.

Cytokines are particularly involved in the immune response as they define the orientation and the strength of the immune response (see section 1.3). Therefore Benavides *et al.* looked for any association occurring between seven microsatellite markers of OAR5 known to harbor some interleukin loci (IL3, IL4 and IL5) [39]. In the end three markers showed significant associations in the Corriedale flock, two of which being also associated to FEC variation in a second Polwarth flock. Interestingly the CSRD2138*A allele was consistently associated with FEC reduction in both flocks thus providing an interesting marker for selection [39]. In a following study using intronic marker located within IL3, IL4 and IL5 loci, Benavides *et al.* showed a significant association for the IL-4 marker in the Corriedale population only but not in the Polwarth flock [40].

4.5 Testing for differential gene expression levels

4.5.1 Introduction

More than a simple marker-trait association that does identify marker in linkage disequilibrium with the causative mutation, difference in gene expression can provide an information about genetic mechanisms, either through absolute or relative quantification [319]. However a difference in one gene expression does not preclude on this gene being the causative as it may also be under the control of another *cis* or *trans* regulator.

Furthermore, it is impossible to detect any post-transcriptional regulatory events. Further care has to be taken while interpreting the results, since false-positive results are likely, due to the large amount of resulting data and validation with other method, e.g. real-time RT-PCR, is necessary. Moreover it can be difficult to eliminate response due to an unknown uncontrolled stimulus if working in field conditions, or to differentiate between cellular types in a tissue sample. This issue might be overcome by using lines of cells that are most likely to be in contact with the pathogen, but this will dramatically simplify the complexity of the system. Lastly resistance may be due to differences in protein structures rather than in their quantity [48].

During the past two decades huge progress has been in molecular genetics that have provided tools for investigating either a few genes with prior knowledge about their function with quantitative real time RT-PCR [230, 231, 220, 561], or providing informations on tens of thousands of genes with microarray [161, 160, 464] or even to the whole transcriptome through RNA sequencing [548].

4.5.2 Investigation of a few genes

The basic approach consists in testing a few genes with prior knowledge of their function that could be related to the trait of interest. Such an approach can be performed through quantitative RT-PCR. In the frame of mining resistance to GIN, the relative quantification is usually employed either for comparing infected and control individuals or comparing breeds or lines with different susceptibility.

Briefly, RNA are collected from a tissue sample and reserve transcripted into cDNA. This cDNA is subsequently exponentially amplified during a limited number of cycles where hybridization with a fluorescent dye occurs. Then PCR reaction enter its retardation phase before reaching a steady state [463]. During the reaction time course, the cycle at which a given

threshold of fluorescence is reached is called the C_t value. This value is the basis for defining gene expression ratios. Critical points of this technique have been discussed by Bustin [75] and required standards for publishing qPCR results have been published [502].

This approach has been widely used to study resistance to GIN as outlined in sections 1.3.

4.5.3 Application of the microarray technology for more exhaustive investigation

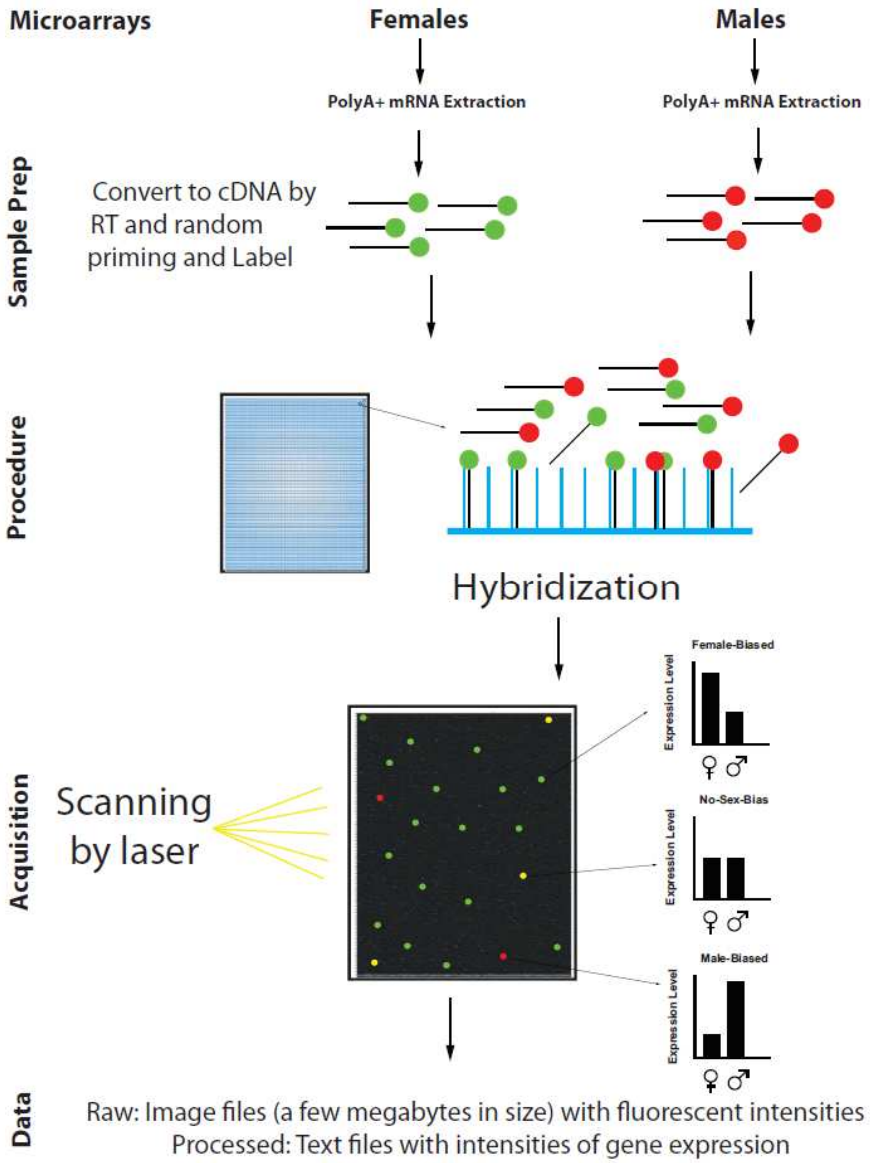
After tissue sampling, RNA are reversed transcripted into cDNA and simultaneously hybridized with fluorescent dye (see figure 4.1). Labeled cDNA are subsequently hybridized on a glass slide where complementary known sequences are bound. After washing, matching cDNA remained on their complementary sequence and the light intensity can be used for assessing gene expression amount [160, 464, 334]. There has been a lot of questioning about working with one or two colors [293, 409]. In a single color microarray experiment both tested groups are labeled with the same colored dye, each sample being hybridized to a unique microarray. On the contrary, in the two-color design two samples are hybridized together on a microarray, each sample being labeled with a proper color. Therefore any higher affinity of one of the two dyes could bias hybridization and the subsequent reading of the fluorescent signal. Nevertheles, the two-color design offers the advantage of reducing technical bias due to multiple manipulations as two samples are considered together. In the end, it has been demonstrated that this question should not be of primary importance while setting up a microarray experiment [409].

The post-genomic era now allows the exhaustive sequencing of every RNA sequences encountered in a particular tissue, called the RNAseq technique that is detailed in a dedicated section (see section 5.7).

From 2005 to present, there has been a few microarray studies in sheep to identify differentially expressed genes between resistant and susceptible animals that are reviewed in this section.

The first microarray study for resistance to GIN in sheep was performed by Diez-Tascon *et al.* who used a bovine specific array comprising more than 10,000 bovine known cDNA sequences [122]. Four resistant and four susceptible Perendale sheep naturally challenged were slaughtered and duodenum samples were taken for RNA isolation and subsequent hybridization on the bovine array. In the end, the authors found more than 100 genes differentially expressed between the two lines of sheep. They observed a higher expression rate of *MHC* class II genes in genetically

Figure 4.1: Principle of microarray data analysis (reproduced from [334])



resistant sheep, but also an up-regulation of a chemokine receptor known to facilitate antigen presentation on MHC class II molecules. This was also associated to an increase in product of genes related to structure and function of the enteric smooth muscle suggesting an adaptation of resistant lambs intestine to expel worms [122].

Another experiment comparing gene expression in hair and wool sheep (mixture of 50% Dorset, 25% Rambouillet, and 25% Finnsheep and of St. Croix and Barbados Blackbelly respectively) challenged with *H. contortus* also used a bovine microarray [327]. The working hypothesis was that hairy sheep are more resistant than breeds bred for wool production. As animals grew under constant challenge, a standard trickle infection with 3,000 larvae during four consecutive weeks followed by an anthelmintic treatment was applied. Then a 10,000 infection dose was given and six individuals were killed at 3 dpi and six more at 27 dpi in each genetic group. Twenty-four sheep remained as control but an accidental infection took place during the study. In the end, hair sheep showed higher expression of genes related to immune cell infiltration, abomasal tissue repair, Th17 response and anticoagulation process that subsequently evolved into up-regulation of genes affecting gut motility and inflammatory response.

Additional work by Rowe with a bovine specific microarray aimed at exploring the evolution of gene expression over time after infection [444]. They used three groups of Merino cross sheep that were challenged three times a week by 500 L3 *H. contortus* ivermectin susceptible and additional L3 *H. contortus* ivermectin-resistant larvae were given three weeks apart. After the second challenge was done with resistant larvae, animals were drenched with ivermectin and subsequently slaughtered after four days. This enabled them to have a look at the impaction of young or adult ivermectin-resistant stages on the abomasal mucosa across time. However no difference between larvae lines could be observed so that the authors bulked all data from the same period together without considering any effect due to the larvae line and to regroup data into three time points at around day 3, 10 and 22. Genes common to three summarisation methods at a p value below 0.01 were considered as well as genes with a function related to veterinary parasitology showing an expression fold change ($p < 0.05$). Their results showed significant change in the expression of genes related to both the innate (intelectin2, TreFoil Factor 3, mucin 5 and calcium chloride channel 1) and the acquired immune response as the trial progressed [444].

Additional findings by Andronicos *et al.* compared four lines of sheep selected for resistance/susceptibility to *T. colubriformis* or *H. contortus* using a bovine microarray [16]. Their

results showed that no common pattern of gene expression could be extracted from the data. However *CXCL10*, that is controlled by IFNγ was up-regulated in the *T. colubriformis* susceptible line.

An ovine specific microarray was designed by Keane *et al.* [266]. Using this microarray, Keane *et al.* [265] found consistent results with previous published results by Diez-Tascon *et al.* [122]. Not only emphasizing the role of the MHC class II locus by demonstrating an 8.4-fold up-regulation of the *DQA1* gene in resistant lambs, the authors also observed an increase in free radical scavenging gene expression [265]. Moreover the same team previously detected that stress response genes were more highly expressed in susceptible lambs before challenge, suggesting that they may have an inappropriate level of expression responsible for perturbation in the innate immune response [266].

In the same way, Knight *et al.* designed an ovine microarray to study differences in gene expression between two groups of animals with different immune status towards *T. circumcincta*, *i.e.* naïve or repeatedly trickled and challenged with a single dose of 50,000 larvae [277]. This investigation also provided a kinetic information on the gene expression pattern from day 2 to day 21. The results showed that immune animals had up-regulated expression of genes associated with cytotoxicity (granulysin and granzymes A, B and H) and pro-inflammatory response as well as mucous associated proteins (calcium-activated chloride channel 1), mast cells associated proteins and tissue remodelling gene. On the contrary, gastric lyzozyme enzymes were down regulated.

Microarray experiments have addressed various questions surrounding GIN infection in sheep, either by comparing impact of the infection by GIN in naive sheep [122, 277], or by comparing sheep with different genetic background put under same infection conditions [266, 265, 327, 16] with [265, 327] or without [122, 16] a kinetic component. More recent studies have tried to dissect the acquired part of the immune process either by comparing the response of naïve and challenged lambs [277] or by comparing lines with different genetic background after a single or three repeated infections [16].

Working on sheep remains costly both in term of work load and money. This matter of fact results in considering less than ten individuals for testing the expression of thousands of genes, thus hampering statistical power. To this regard, some studies reported differential gene expression without any correction for multiple testing [444, 327] hence making it difficult to ascertain true up- or down- regulation. For instance, McKinnon *et al.* estimated that 180

false positive signals were expected while reporting 60 transcripts with differential expression [327]. The variation in biological questions, corresponding experimental designs and statistical treatments of produced data are certainly responsible for the lack of unity in the functional genes being involved.

Chapter 5

Looking for the regions underlying resistance to GIN with molecular markers: methodology and applications

Most of the traits of economic value, among which are traits related to resistance to parasites, are quantitative, *i.e.* continuously distributed [151]. Classical quantitative genetics considers that phenotypes are under the control of an infinite and unlinked number of loci, each contributing with an infinitesimal effect [151]. This theory thus considers the genome as a black box [151].

Even if it is not always necessary to dissect the genetic structure of traits at a molecular scale, considering a finite loci model with a few Quantitative Trait Loci (hereafter denoted QTL) controlling the trait of interest, can lead to useful applications [325]. Indeed, in the frame of breeding programs for resistance to disease, this could help identifying genes controlling susceptibility to disease and improving selection programs. From a theoretical standpoint this can also make theory more realistic [151] as mammalian genome is known to carry a limited number of genes (thought to be around 30,000 genes in the case of the human [148]).

Some knowledge on phenotypes of interest can help directly targeting functional candidate genes and test if those genes are particularly involved in the phenotype of interest (as described in previous chapter 4). However, this candidate gene approach is hampered by two major problems. Firstly, quantitative traits are often under the control of many genes hence making it

difficult to target a restricted number. Secondly, in the case of disease resistance, this approach is limited by the *a priori* knowledge of immune mechanisms at stake during host infection.

A more indirect approach is to map QTL, *i.e.* to identify regions of the genome that are particularly associated with phenotypic variation in resistance traits. This chapter aims at reminding the basic principles of QTL mapping as well as the associated requirements, with an emphasis put on properties of linkage disequilibrium in sheep populations. Another section will detail the different methodologies available whereas an extensive literature review of QTL mapping applied to resistance to GIN in sheep will be done in the last section of this chapter.

5.1 Definitions and basic principles of QTL detection

QTL stands for "Quantitative Trait Loci". It refers to a region of the genome affecting a quantitative trait in a non negligible way. This concept does not forecast the architecture of the genetic region: it may be a single gene with a moderate effect or a cluster of genes with more modest effects that aggregate to determine a significant proportion of the genetic variation. So far it has been extremely difficult to determine underlying causative mutation, *i.e.* QT Nucleotide (denoted QTN) as exemplified by the extreme ratio between the four identified QTN reported and confirmed in livestock [552] and the 15,014 reported QTL stored in the QTL database [240] (database accessed the 15^{th} of march 2012). However this huge amount of unraveled QTL tends to confirm that quantitative traits are controlled by many genes of small effect and very few genes of large effect [216, 92, 552].

First attempts of mapping QTL were done at the beginning of the XX^{th} century when Payne demonstrated that scutellar bristle number was found to be influenced by X chromosome in Drosophila [410] in [325]. Subsequent pioneer QTL mapping studies have been reviewed by Lynch & Walsh [325] and will not be detailed in this manuscript. All these early attempts aimed at using objective markers that would score variation at a molecular level in order to explain the observed phenotypic variation. Requirements for a successful QTL mapping thus come straightforward: variation in the considered phenotype and polymorphic genetic markers linked to the QTL, whose relative positions are known and that adequately cover the genome [151].

The assessment of whether a chromosomal segment carries a QTL consists in two steps: clustering the phenotyped population into various classes based on the allelic version the individuals

carry, and performing a statistical test to determine whether any significant difference can be observed between phenotypic means of each allelic group [176]. Over a certain threshold this test statistic will determine the presence of a QTL on the investigated region and an associated confidence interval of the true position will also be derived. It thus follows that a QTL is rather a statistical concept that can only be determined by the unraveling of the underlying QTN. Without this crucial confirmation step, a QTL will remain a probability distributed over a more or less wide chromosomal region.

5.2 Genetical toolkit for QTL mapping

5.2.1 Molecular information

5.2.1.1 Molecular markers and their properties [325, 535]

Following the initial works of Payne and other chromosomal assays performed by Thoday, first molecular markers were generated in the late 1960s, being either allozymes or protein variants. DNA-based molecular markers came twenty years after and their basic properties are summarized in table 5.1.

Among these Restriction Fragment Length polymorphisms (RFLP) are among the most simple. DNA is firstly digested by restriction enzymes that cut DNA strand at precise sequence called restriction site. The DNA fragments subsequently hybridize with specific labeled DNA probes that are complementary of specific regions of the genome. Heterozygous and homozygous can be differentiated through this method.

Randomly Amplified Polymorphic DNA (RAPD) markers are revealed by the use of random short primer sequence and PCR. If two primer binding sites are not too far apart, the PCR is successfull and a DNA fragment is amplified. If one site is missing no DNA amplification will occur and no fragments are generated for the region. This type of marker is dominant as fragment will not be produced in the homozygous missing the specific site solely. These markers are easy to use as only standard oligonucleotides are required hence also making them very affordable. But they are generated at random and they lack reproducibility as they highly depend on the PCR conditions.

A close type of markers are the Amplified Fragment Length Polymorphisms (AFLPs). In this case, DNA is firstly digested by restriction enzymes and fragments are subsequently amplified by PCR and separated through electrophoretic separation [542]. This is a mixture between RFLP

and RAPD markers.

Table 5.1: Information content and requirements of molecular markers

		Information content		Requirements					Characteristics	
Marker name	Dominant	Co-dominant	Restriction	PCR	Specific primers	Gel	Development effort	Genotyping effort	Reproducibility	Accuracy
RFLP	-	+	+	-	-	+	High	High	High	Very high
PCR-RFLP	-	+	+	+	+	+	High	Medium	High	Very high
RAPD	+	-	-	+	-	+	Very low	Very low	Low	Very low
AFLP	+	(+)	+	+	-	+	Low	Very low	High	Medium
SSCP	-	+	-	+	+	+	Medium	Medium	Medium	Medium
Microsatellite	-	-	-	+	+	+	High	Low	High	High
SNP	-	+	-	+	+	+/-	High	Variable	High	Very High

Adapted from Vignal et al. [535]

RAPD or AFLP, also known as fingerprinting techniques, are particularly devoted to species for which only a limited number of microsatellites have been developped, the AFLP technique being of choice for inter-laboratories results sharing. However these two techniques generate dominant markers only thus requiring more markers to achieve the same power as microsatellites. As their location is at random and unknown, large datasets may lead to the use of linked markers which is contradictory to the usual assumption.

Another type of marker, called microsatellite, shows variation in the length of short sequences of tandemly repeated DNA. They are highly polymorphic and are specifically dedicated for analysis in outbred populations. Therefore they have become markers of choice in the end of 1990s, hence replacing RFLP, as a simple PCR followed by allele sizing on polyacrylamide gel gives access to the genotype.

Over the last ten years, Single Nucleotide Polymorphisms (SNP) have known an increasing popularity relying more on the need of very high densities of markers than on the genetic information provided. These markers are single base changes in a DNA sequence. Such variations in DNA are either due to mutations occurring at a very low rate (estimated to range between 10^{-9} and 5.10^{-9}) or through transition, that are usually biased towards transition (purine-purine or pyrimidine-pyrimidine) rather than transversion. These two factors determine the fact that SNP are usually biallelic markers.

Due to their biallelic property, SNP provide less information than microsatellites which in turn requires more SNP than microsatellites to reach same performances for detection of admixture, estimating inbreeding coefficient or study relationship between individuals. However their price is much lower than for microsatellites (see table 5.2). In addition SNP will be of particular interest for fine mapping of QTL as they constitute the most common polymorphism and as their presence between and within genes offer the possibility to better target genes of interest, including exons and promoters. They are also highly reproducible from lab to lab.

Table 5.2: Comparison of microsatellites and SNP markers

	Microsatellites	**SNP**
No. markers	150 - 200	50,000 - 60,000
No. polymorphic alleles	2-20	2
Marker density	0.5 /cM	30 /cM
Localization	between genes	within and between genes
Price (in euro)	200-300	100-150

5.2.1.2 Genetic map [325]

Genetic maps show both the relative ordering of markers along the linkage group and the relative distance between them. The recombination frequency between two loci roughly depends on their physical distance. However it is not a suitable as a genetic distance as a distance should be additive. Indeed considering three different ordered loci, A, B and C with recombination frequencies c_{AB}, c_{BC}, c_{AC}, the recombination frequency between A and C can be expressed as follows:

$$c_{AC} = c_{AB} + c_{BC} - 2(1-\delta)c_{AB}c_{BC}$$

, with δ the interference parameter that is 0 for independent crossing-overs or 1 if a crossover in one region suppresses crossovers in adjacent regions

Two functions of genetic distance have been proposed. The Haldane's function considers no interference :

$$m = -\frac{ln(1-2c)}{2}$$

, where m is the mapping function and c, the observed recombination frequency.

Kosambi function does allow for small interferences :

$$m = \frac{1}{4}ln(\frac{1+2c)}{1-2c})$$

Both approaches reach the same values for extreme situation, *i.e.* m = c for small c and m = 1/2 for large c.

No universal relationship can be proposed between physical and genetic distance. Indeed crossing-overs occurrence varies a lot along the genome with some recombination hot-spots being described [544].

5.2.2 Linkage disequilibrium

5.2.2.1 Definition and usual measures: *D'*,r^2

Any departure in frequencies of gametic types from what would be expected from allelic frequencies is linkage disequilibrium (LD) or gametic phase disequilibrium. The linkage disequilibrium

D can also be derived as [310, 232]:

$$D = f_{11} \times f_{22} - f_{12} \times f_{21}$$

where f_{11}, f_{22}, f_{12} and f_{21} are the respective actual frequencies of the A1B1, A2B2, A1B2 and A2B1 genotypes.

Hedrick [219] gives properties that can be expected of an ideal estimate of LD. It should have a simple biological interpretation and should be mathematically related to populational genetics factors (e.g. genetic drift, selection etc.). In addition statistical tests should be available to estimate it and comparison between pairs of loci or across populations should be feasible [219]. Numerous measures of LD based on D have been proposed (see table 5.3) as D is more an intuitive concept than a useful numerical tool to assess LD since it highly depends on allelic frequencies[19].

Development of LD measures is a rather large and complex topic and great controversy still exists about the best measure of the extent of linkage disequilibrium [474]. However two measures, D' and r^2 described hereafter are commonly used for biallelic markers [310, 233, 19, 159, 579]. Based on reviews [219, 121, 254] we hereafter summarize their main properties and shortcomings with a special emphasis on D' and r^2.

D' was firstly introduced by Lewontin [310] before being extended to a multi-allelic model by Hedrick [219]. This normalized quantity is obtained by dividing D_{ij} by its maximum possible value given the allele frequencies p_i and q_j at the two loci i and j respectively. It is written as follows :

$$D' = \sum_{i=1}^{k} \sum_{j=1}^{l} p_i q_j |D'_{ij}|$$

with $D'_{ij} = \frac{D_{ij}}{D_{max}}$ and $\begin{cases} D_{max} = min[p_i q_j, (1-p_i)(1-q_j)] & when D_{ij} < 0 \\ D_{max} = min[p_i(1-q_j), (1-p_i)q_j] & when D_{ij} > 0 \end{cases}$

The range of D' is normalized to vary between -1 and +1 independently of the p_i [311]. Values below 1 will indicate that the complete ancestral LD has been disrupted but such values do not have clear interpretation [19].

The second often-used measure of linkage disequilibrium for biallelic markers is r^2 which was subsequently devised in 1968 by Hill & Robertson [233]:

$$r_{ij}^2 = \frac{D_{ij}^2}{p_i(1-p_i)q_j(1-q_j)}$$

This coefficient ranges between 0 and 1 and its value has a rather intuitive biological interpretation since its square root is the correlation coefficient in allelic state between alleles in the same gamete [213]. A value of 1 is obtained if and only if no recombination occurred between the markers and if they have the same frequency. It is synonymous of *perfect* LD [19]. In addition for a population of size N, Nr^2 follows a χ^2 square distribution [233] which in turns is of particular interest while testing the null hypothesis of independence between pairs of markers [422].

Linkage disequilibrium reflects population history. To this regard, *D'* and r^2 exhibit different but complementary properties, so that they both have been used to estimate LD in livestock [152, 196, 496, 353, 354, 270, 174, 511, 8, 209, 3, 221] or in human populations [19, 504]. For weak linkage disequilibrium both values will obviously tend to 0 but on the way up to 1, r^2 can take on any value between 0 and D'^2 as it depends not only on D but also on allele frequencies [213]. Supposing a population where the different haplotypes between two loci are AB, ab and aB, the a and b alleles resulting from two consecutive mutations. In that case, $D' = 1$ as $D = P_{AB}P_{ab} - P_{Ab}P_{aB} = D_{max}$. On the other hand, r^2 will depend on when the B to b mutation occurred: if the mutation is ancient, then the latest created haplotype (ab) will have spread in the population so that the resulting correlation between a and b will be high. On the contrary a very recent mutation will result in a low value of r^2. Therefore *D'* only measures recombinational history whereas r^2 summarizes both recombinations and when and where mutations have occurred [159, 213].

Even if complementary these two statistics both "*obscure the direction of linkage disequilibrium*" since the original sign of D is lost [213]. Another shortcoming is that neither r^2 nor *D'* perform well with small sample size, estimation of LD with *D'* being more erratic [19, 159, 579]. Due to haplotypes of low to zero frequencies, *D'* tends to inflate LD estimates even between markers almost in equilibrium [579].

As introduced before other measures of LD have been devised and have been used in human case/control studies but have limited applications for animal population (summarized in table 5.3 from [573]).

Table 5.3: Measures of LD between 2 biallelic loci

LD measure	Expression as a function of D	Reference
D	$D = p_{AB}p_{ab} - p_{Ab}p_{aB}$	[151]
r^2	$\frac{D^2}{p_A p_B p_a p_b}$	[233]
D'	$\frac{D}{D_{max}}$	[310]
δ	$\frac{D}{p_B p_{ab}}$	[41]
d	$\frac{D}{p_B p_b}$	[385]
Q	$\frac{D}{p_{AB}p_{ab}+p_{Ab}p_{aB}}$	[575]

Key: p_{ij} is the probability of carrying i and j at first and second locus respectively, and p_i the probability of having allele i

5.2.2.2 Factors of variation of LD in livestock populations

LD reflects the history of a given population as haplotypes present in this population shared the same ancestry [213]. To this regard every factor that usually affect animal populations, *e.g.* migration, mutation, selection, will affect the two successive sampling events described by Weir and Hill [550] that result in variation in LD:

- the sampling of gametes conditioned by the mating system and the effective population size, N_e (that is number of individuals that would give the same actual population inbreeding if they were bred in the manner of the idealized population [151])
- the sampling of individuals directly depending on demographical considerations

Mutation is the base material to create LD. When a mutation appears on a given haplotype it is specific of this haplotype and the alleles it carries (see last subsection 5.2.2.1). However mutation is a rare phenomenon, occurring at a rate of about 10^{-5} to 10^{-6} per generation [151]. Therefore mutation alone will have a very low impact in changing gene frequencies in a population except for some markers known to have high mutation rates hence resulting in little to null LD with nearby markers [19].

But the selection of this mutation in the subsequent generations, either randomly occurring through genetic drift or being achieved through genetic selection, will result in haplotypes being lost and hence determine various levels of LD [233, 550, 213]. Selection is hence of particular interest in case of livestock populations and has been extensively studied [310, 155, 383, 510]. Selection affects a population in the same way as genetic drift does, as it tends to reduce the number of haplotypes. As such, it is clear that it will either increase the extent of LD over

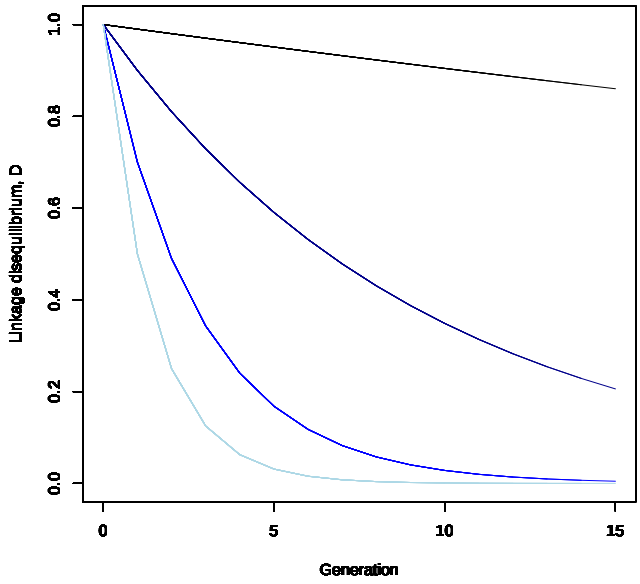
genetic distance, or at least will slow down its decay.

On the contrary, recombinations breaks down chromosome segments. This removal occurs in case of random mating at a rate of $1 - c$ per generation. D thus naturally tends towards 0. It can be derived (see [151] for an extensive derivation) that after t generations:

$$D_t = D_0(1 - c)^t$$

where D_t is the LD at generation t, D_0 is the initial value of LD and c the recombination rate between the two considered loci.

Figure 5.1: LD decay as a function of the generation time for different recombination rates



A graphical representation of this formula (figure 5.1) for different recombination rates c shows that LD decay is slower for tightly linked markers. Indeed some very close markers on a chromosome are bound to be passed together to the next generation. However one has to keep in mind that LD is not a measure of physical linkage as even loci that are far away can exhibit some high level of LD and conversely [422, 19].

It thus follows that LD between close markers reflects ancient recombination processes as well as drift. On the contrary, LD between markers far apart translates the more recent history of a population. To this regard, it is worth noting that the genetic markers considered for LD analysis are common variants that hence tend to be older than random polymorphisms and that thus will reflect older events [422].

Other demographical parameters can affect LD. Admixture,*i.e.* the introduction of genes from one distinct population into another, can create spurious LD [384]. Indeed the mixing of two populations will affect the allelic frequencies thus modifying LD patterns; this is particularly true in dairy cattle with the globalization of semen [152]. In subsequent generations and without any selection occurring, this will be rapidly dissipated between unlinked markers whereas recombination will decrease LD at a slower rate between linked markers [384]. In the same way a structured population in which a particular allele only segregates in a given subpopulation will tend to give rise to spurious LD, this allele being often associated to other markers specific of this subpopulation [423, 425, 475].

5.2.2.3 Linkage disequilibrium in animal populations with a special emphasis on sheep

Given the described properties of LD, following requirements can be made for estimating LD in a population:

- "unrelated" animals
- sufficient sample size
- choosing the appropriate neutral genetic markers

First estimation of LD in an animal population was done by Farnir et al. in 2000 for dairy cattle [152]. So far, many domestic animals, *i.e.* cattle [152, 174, 511], sheep [353, 354, 270], pigs [8, 209, 26], chicken [3, 221], dogs [196, 496], horses [99, 509] and even the rainbow trout [434], have been characterised for LD extent. Most of these studies found that:

- LD decreases rapidly with physical distance
- Some non negligible level of LD could be observed even between non syntenic loci

A few studies have been conducted in sheep with microsatellites [353, 354] or SNP markers [270, 274, 372]. Using microsatellites McRae *et al.* (2002) and Meadows *et al.* (2008) found consistent high level of LD between syntenic markers of about 0.2 associated to a rapid decay of LD with marker distances especially in Merino and Coopworth breeds [353, 354]. In some breeds, *e.g.* White Faced suffolk and Poll Dorset, some intermediate rates of decay were observed and interpreted as a result of recent bottleneck occurring during breeds formations [354]. This

relatively high rate of LD decay with marker distance was retrieved by Kemper *et al.* (2011) [270] who estimated average r^2 in the same breeds as Meadows *et al.* [354], namely Merino, Border Leicester cross Merino, Poll Dorset and white-faced suffolk. Mean LD coefficients ranged between 0.12 for Merinos and 0.19 for Poll Dorset, the difference between the two translating differences in effective population size. Additional work performed by Miller in wild breeds of America (thinhorned breed from Alaska, bighorned Mountain and Wyoming breeds) showed similar level of LD and they found an average half length LD, *i.e.* the genomic distance by which LD dropped by 50% of its maximal value [432], of 4.6 Mbp being in the same range of what has already been found in livestock [372].Interestingly, Kijas *et al.* estimated the effective population sizes of 74 sheep breeds genotyped for 50K SNP and found that 75% of sheep breeds have N_e of 300 individuals or more that contrasts with much more reduced N_e estimated in cattle (150 individuals) [180].

5.3 Statistical toolkit for QTL mapping [325]

Several types of analyses are available for QTL mapping that all exploit existing linkage disequilibrium either historical LD at the population level (association analysis) or more recent LD pattern through the observations of recombinant animals within multiple families (linkage analysis). Meuwissen and Goddard [365] summarized QTL methodology in three steps being the calculation of a probability G_{ij} that two individuals carry chromosomes being identical by descent (either within or across families), the comparison of the phenotypes according to G_{ij} and finally finding the position that maximizes the likelihood of the phenotypes given G_{ij}. Recent developments have also aimed at combining both knowledge, *i.e.* ancient and recent LD, to benefit the advantages of both analyses. This section summarizes the main characteristics of each analysis.

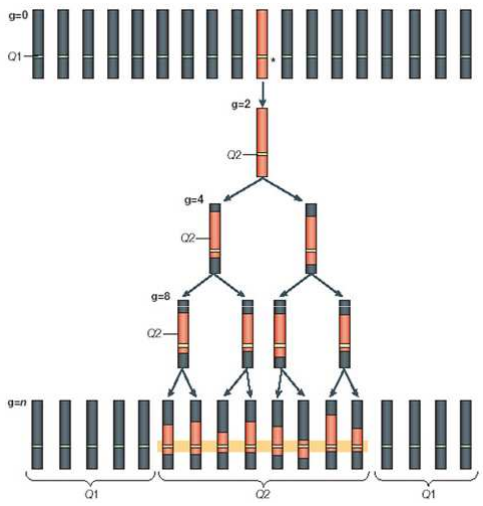
5.3.1 Linkage Disequilibrium Analysis (LDA) or Association analysis

This analysis is the rough application of the QTL mapping principle that is looking for any significant relationship between a marker genotype and a recorded trait across a given population. Therefore no tracing of chromosome inheritance is performed as it is done in the so-called "within family analysis" (see next paragraph 5.3.2) [285].

LDA takes part of the recombination history of the population since its creation so that

QTL position can be pinpointed as showed in figure 5.2 [14]. Assuming a beneficial mutation occurs at the generation n on a given chromosome, the recombination events in the subsequent generations are expected to break down the initial chromosomal segment and to select for it. In the end, only a tiny fragment of the original chromosome remains in the actual population. Assuming a marker is close to this fragment, the QTL position can be pinpointed to the area in linkage disequilibrium with the QTL.

Figure 5.2: Schematic representation of the recombination history in a given population (reproduced from Andersson & Georges [14])



It thus follows that the power of the analysis, that is the probability of correctly rejecting the null hypothesis (no effect of the marker on the trait of interest) when a true QTL exists, is directly conditioned by the size of the QTL effect, QTL frequency, and the linkage disequilibrium between the QTL and the associated marker. Considering the latter issue, it has been estimated that to

achieve the same power as if the QTL genotype was known, the sample size should be increased by a factor of $\frac{1}{r^2}$, r^2 being the LD value between the marker and the QTL [422]. Therefore, the required density of markers made it impossible to effectively use LDA for mapping purpose until the release of DNA SNP chip, as for instance, a "useful" range of LD was predicted to be less than 3 Kb in the human population [285, 286]. However, provided that the QTL has a beneficial effect, it is probable that other markers in its vicinity have been selected for so that LD should extend over larger region, thus making it possible to map the QTL through indirect association [234, 580, 287]. A study performed by MacLeod *et al.* (2008) investigated the particular case of livestock populations, where family structure is important [328]. They estimated a 37% power of detecting a QTL explaining 5% of the phenotypic variance for a flock of 365 cows genotyped with 10,000 SNPs and they also reported strong correlations between SNP associated statistics and their r^2 value with the QTL [328]. Their results nuanced the optimistic findings of Zhao *et al.* [580] that estimated that given the population structure around 4,000 to 6,000 SNPs would be necessary for whole genome association study. Working with the proportion of the phenotypic variance explained by the QTL (σ_p) summarized both the QTL effect (α) and its allelic frequency (p) given that [328]:

$$\sigma_p = 2p(1-p)\alpha^2$$

This relationship recalls that the allelic frequency of the QTL is also an important parameter conditioning the QTL identification, as the rarer the QTL, the bigger its effect should to explain a moderate proportion of the observed variation.

One way to overcome the problem of low level of LD between a QTL and genotyped markers is to consider haplotypes of markers, especially in the case of recent mutations [29, 345]. This strategy can provide additional knowledge about existing *cis* acting regulatory region close to the causal gene [29].

In addition sample size and the minor allele frequency of the associated variant (that conditions the number of individuals available for estimating an effect) also impact the power of detection [580, 328].

In addition to this problem, the structure of the population can greatly affect the outcome of LDA [425, 29, 234]. Indeed the pattern of LD is bound to vary according to the considered population. For instance the recombination pattern observed in figure 5.2 may not be the same

in another population. If the population under study is a mixture of several populations with different histories, then association signals for a given phenotype might translate a subpopulation characteristic different from the phenotype of interest. To illustrate this point, Astle & Balding took the example of a large sample of individuals drawn from the UK population [22]. In this sample many markers are expected to be associated to the phenotype "speaks Welsh" as their allelic frequencies will be specific of Welsh people due to population history. However these markers will not be causative of the speaking Welsh phenotype. In their study, MacLeod *et al.* found that not taking into account the genetic structure of the population while performing LDA resulted in increasing the number of false positive signals [328].

Pritchard *et al.* developed a method of clustering individuals according to their allelic frequencies at multiple loci [424]. Information on the population structure is subsequently used to perform LDA within subpopulations [425]. Another approach proposed by Price (2006) for case/control study was to apply a principal component analysis (PCA) to genotype data in order to infer population ancestry and to correct both genotypes and phenotypes by the values obtained along each axis of the PCA before testing for association [420]. This had also been suggested by other authors [578]. This analysis was subsequently outperformed by the EMMAX (Efficient Mixed-Model Association eXpedited) model that is inspired from animal genetic models and that infers the cryptic relatedness of individuals through the computation of a relationship matrix based on dense set of markers [257]. An extensive review of methods taking population structure into account has been done by Astle and Balding [22]. This issue of population structure is rather a human genetics problem as animal population structure has been recorded into deep pedigrees whose knowledge corrects for structure.

5.3.2 Linkage Analysis (LA)

5.3.2.1 The simple case of inbred lines

As already outlined before, QTL mapping relies on the association between genetic markers and putative QTL. Until recently, available markers density was not sufficient to exploit LD existing across animal populations so that particular experimental crosses were implemented.

Taking the example of inbred lines, F_1 reproducers are created by crossing individuals from two parental lines (F_{01} and F_{02}), and F1 individuals are either back-crossed to one parental line (F_1 x F_{01} or F_1 x F_{02}) or mated together to produce F_2 progenies (F_1 x F_1). Through crossing, LD is generated while a limited number of recombination events decreased it, still making it

possible to exploit it with a sparse genetic map. In addition, for each genotyped marker it is possible to determine its inheritance as the pedigree was clearly known.

It follows that the more meiosis occur, the more recombinant can be observed and the more precise will be LA mapping. Darvasi *et al.* (1993) demonstrated that population size and QTL effect were the two critical factors affecting the accuracy of LA. Increasing the marker density to an infinite number was predicted to increase the accuracy up to the 95% confidence interval (interval that contains the QTL with a 95% probability) but not beyond [106].

In the end, within a family, individuals can be clustered into two (back-cross population) or three (F2 population) groups according to their genotype at each considered genetic marker. In the case of a back-cross (BC) population, half of the individuals are homozygotes, the other half being heterozygotes. At a given marker position, any significant difference between the two groups would indicate the presence of a QTL.

5.3.2.2 Interval mapping

The use of single marker information greatly hampers QTL detection as it implicitly assume that both genetic marker and QTL are confounded which is far from achieved with sparse genetic map. Lander & Botstein (1989) stated that for a recombination fraction θ, the effect of the QTL is underestimated by a factor of $(1 - 2\theta)$ [292]. In addition no precise QTL location can be inferred, as the observed probability that a QTL is actually in linkage with the considered marker can be due either to a big QTL far away from the marker or a tightly linked QTL with a weak effect [292, 325].

To overcome the related issues of simple marker trait association with sparse map, Lander & Botstein proposed an interval mapping approach using likelihood method [292]. The idea is to exploit the known genetic distances between available markers to compute the probability that a QTL lies in between them. This probability is modeled as a likelihood function that a QTL is linked to the marker, knowing the marker genotypes and the observed phenotypes [325].

For a BC inbred population this gives the following expression:

$$L = \prod_{i=1}^{N} \sum_{j=1}^{2} Pr(Q_j|A_iB_i) f(y_i|m_j, \sigma^2)$$

where,

$$f(y_i/m_j, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} exp[\frac{-(y_i - m_j^2)}{2\sigma^2}]$$

Q_j is the genotype at the QTL, with genotype A_i and B_i at markers A and B for individual i; m_j is the effect of QTL genotype, σ^2 is the residual variance and N the number of BC individuals.

In the end a lod-score (LOD) is computed as the likelihood of the alternative hypothesis over the likelihood under the null hypothesis [292]. This LOD is equivalent to a likelihood ratio (LRT) that is usually implemented in animal genetics. The relationship between the two statistics is given by [325] :

$$LOD = \frac{LRT}{2ln10} \approx \frac{LRT}{4.61}$$

Linear regression methods, easier to compute, were proposed for F2 [207] or BC [339]. The idea is to express regression coefficients as a function of the unknown QTL parameters [207, 325] using the following modeling of the phenotype y_i :

$$y_i = \sum_{j=1}^{2} m_j Pr(Q_j|A_iB_i) + e_i$$

For a QTL with effect being $\mu + a$, $\mu + d$ for the QTL genotypes QQ, Qq respectively, the authors derived expected mean performance of BC individual for each genotype. The a and d values corresponding to a putative QTL can thus be calculated within marker interval for each tested position. In the end, the regression and residual sums of squares can be calculated and the position that minimizes the residual mean square is the most probable QTL position [207]. This was shown to be very similar to likelihood estimates especially for QTL of small to moderate effects [207, 430]. Kao [258] also found this type of results but underlined the poor performances of linear regression for detecting closely linked QTL.

5.3.2.3 Back to reality : the case of outbred populations

As to be opposed to plants or mice, livestock populations are outbred. This has two major practical consequences for QTL mapping [325]:

- Marker and QTL are usually not in complete LD, hence resulting in different linkage phases within a given population and non-informative families, *i.e.* parents homozygous for the QTL allele or for the associated marker or both

- QTL effect is expressed as a variance and not as the average value of QTL genotype like for inbred lines, hence introducing bias in its estimation
- Likelihood computation is complicated due to the need to consider all possible genotypes of measured progenies

Three generation crosses, *i.e.* BC or F2 populations can also be used with outbred lines and the QTL detection analysis can be performed assuming QTL alleles are fixed within each each line [278]. Likelihood is computationally demanding [325] and the need to consider more possible genotypes in outbred populations makes QTL mapping difficult with such methods [325]. Therefore extension of the linear regression model have firstly been derived for inbred lines has been proposed for F2 by Haley & *al.* [208]. Once again phenotypic values are regressed upon additive and dominance coefficients of the QTL at a given position knowing marker genotype. For each individual and each tested position, marker genotypes are known hence making it possible to determine line origin (line 1 or 2) of the considered chromosomic segment, *i.e.* the QTL, with a probability expressed as a function of the recombination frequencies between markers [208]. In the end, QTL-associated genetic coefficients (additive effect a and dominance effect d) can be estimated as :

$$a = prob(\omega_{11}|P) - prob(\omega_{22}|P)$$

$$d = prob(\omega_{12}|P) + prob(\omega_{21}|P)$$

, where prob(ω_i — P) is the probability of line origin combination i for a QTL at a given position conditional on the observed marker genotypes in this progeny and its ascendants [208].

It seems linear regression and likelihood approaches perform the same in terms of power of detection [208].

Subsequently to this work by Haley & *al.* (1996), simplifications have been proposed by Elsen & *al.* (1999) that finally provide a relatively easy framework for QTL detection using likelihood in livestock populations [141].

5.3.3 Joint linkage and association analysis (LDLA)

As outlined in the two previous subsections, both LA and LDA are hampered by disadvantages being respectively a lack of precision or the handling of false positive signals. Joint linkage and

association methods have been derived that benefit both the robustness of the linkage analysis and the precision of the association analysis, [366, 153, 307].

One of the first examples of combining the two analyses was reported by Riquet *et al.* in a two-step approach [436]. An initial QTL mapping using LA identified seven heterozygous "Qq" sires. Additional markers were developed within the QTL region and authors looked for identical by descent haplotypes common to each of the seven sires thus pinpointing the QTL position to 5 cM interval [436].

Subsequently Meuwissen and Goddard proposed a one step methodology to fine map QTL using both linkage and linkage disequilibrium information [366]. Their analysis estimates variance components for the QTL associated to markers, the effect of background genes (estimated from the genetic relationship matrix built with the pedigree) and an error term. In the end the twinning rate they investigated was modeled as :

$$y = \mu\mathbf{1} + \mathbf{Zh} + \mathbf{u} + e$$

where μ is the overall mean, $\mathbf{h}$, $\mathbf{u}$,$\mathbf{e}$ are the vectors of random haplotypes, random polygenic and random sampling errors and $\mathbf{Z}$ is an incidence matrix.

For the estimation of the genotype probability, two cases can be distinguished as the common founder can either be within the pedigree or unknown. In the first case, the analysis is similar to a classical LA in which founder allele is traced and the probability of inheritance is estimated weighted by recombination rate. In the second case, the probability that the haplotype is inherited from a common ancestor is estimated as reported in [362]. Finally a full matrix of pairwise IBD probabilities between haplotypes is computed.

Farnir and colleagues [153] simultaneously (paper published 11 days after the one by Meuwissen & Goddard) implemented a similar approach to the QTL region affecting milk yield and investigated by Riquet *et al.* [436]. The probability that a QTL was present at a given position was modeled as the following likelihood [153]:

$$L_{NPeds} = \prod_{m=1}^{M} \sum_{k=1}^{A} f_k \prod_{i=1}^{N} \begin{bmatrix} P(Sire_i = Q_A Q_B / m_i) \times P(Ped_i / Sire_i = Q_A Q_B) & + \\ P(Sire_i = q_A q_B | m_i) \times P(Ped_i | Sire_i = q_A q_B) & + \\ P(Sire_i = Q_A q_B | m_i) \times P(Ped_i | Sire_i = Q_A q_B) & + \\ P(Sire_i = q_A Q_B | m_i) \times P(Ped_i | Sire_i = q_A Q_B) & \end{bmatrix}$$

where L_{NPeds} is the likelihood over the M markers, the k alleles and N half sib progenies that there is a QTL, $P(Sire_i = q_A q_B / m_i)$ is the probability that the sire i carries the q_A and q_B alleles knowing its genotype at the m_i marker and $P(Ped_i / Sire_i = Q_A Q_B)$ is the probability contributed by the i^{th} progeny to the likelihood knowing its $Sire_i$ genotype $Q_A Q_B$.

In this model the probabilities of the QTL genotype knowing the marker genotypes are derived as a function of the recombination rate between the marker and the QTL loci, and the frequency of the marker alleles. In the absence of linkage disequilibrium this probability can be simplified to the one computed in a linkage analysis [153], so that linkage and linkage disequilibrium are exploited simultaneously.

Recently Legarra & Fernando (2009) proposed a linear regression model for a joint LA and LDA analysis [307]. Within sire QTL effects are added to the sire haplotypes effects of the LD-decay model in order to account for a possible between-sire variability of the QTL effect beyond that reflected by the haplotype [307]. This model requires less computational effort than the IBD method proposed by Meuwissen & Goddard (2002) [366].

Meuwissen & Goddard (2002) could refine the QTL position to a narrow region whereas both LA and LDA approaches resulted in more numerous and wider QTL region [366]. The same gain of precision was obtained by Druet *et al.* (2008) that mined a QTL region affecting fertility in dairy cattle and finally decomposed a 40 Mbp region into 8 narrow intervals [133].

The method proposed by Meuwissen & Goddard assumes that linkage disequilibrium phase is conserved among the studied population. However studies on LD pattern in dairy cows populations showed that correlations between markers genotypes were not always conserved between breeds [114]. Taking this into account, Uleberg *et al.* tuned the original IBD model to apply a LDLA method on a commercial pig cross of different lines [517]. The tuning consisted in setting the IBD probabilities between founders from different breeds to zero, hence being considered as having no common ancestry [517]. However the results produced with this breed model did not differ from the original one suggesting that considered lines were more similar than expected [517]. Still this point of admixed population remains challenging while using LD for mapping purposes.

5.3.4 QTL mapping aims at precisely mapping a true QTL

5.3.4.1 Fitting the right threshold to avoid false positive

QTL analyses provide test statistics associated to various positions along the genome. A threshold is therefore required to identify the most probable regions harboring genes of interest [325, 291]. As outlined in the previous subsections, a very high number of tests are to be performed for QTL mapping purpose hence multiplying the associated risk of false positive associations occurring. For each statistical test performed, a threshold needs to be defined to differentiate between significant signals and other false positive associations that arise just by chance. This is particularly true in the case of LDA as a test is performed for each SNP and experiments usually employ at least 50,000 and up to 1,000,000 of SNPs in human studies.

However the distribution of the tests statistics is unknown. The most straightforward approach is to apply a Bonferroni correction, that supposes independence between tests and fits a significance level of $\frac{\alpha}{n}$ [325]. However this threshold is usually highly conservative and it does not fit the biological reality of physical chromosomal reality and existing LD pattern as every test are considered independent. Another approach can be to quantify the False Discovery Rate (FDR), that is the proportion of false positive among positive results [491, 157, 29].

While proposing an interval mapping approach, Lander & Botstein also provided an approximate threshold for defining significant QTL as :

$$T_\alpha = (2log(10))t_\alpha$$

where t_α solves the equation $\alpha = (C + 2Gt_\alpha)\chi^2(t_\alpha)$ for C chromosomes and a genetic map of G Morgans.

Churchill & Doerge have proposed to estimate an empirical threshold for each mapping experiment through appropriate reshuffling of the data [90]. Observed phenotypes are randomly shuffled over individuals genotypic data hence creating a new dataset. The genetic information is therefore conserved while phenotypes are randomly assigned. Threshold value therefore depends on the marker density, the less markers, the lower threshold, but not on the presence or absence of a QTL [90, 292]. Resampling a 1,000 or 10,000 times thus provides an empirical distribution of the test under the null hypothesis. Genome-wise threshold are subsequently obtained by applying a Bonferroni correction based on the number of chromosomes under investigation. Authors recommended to perform at least 1,000 permutations for a 5% significance level [325,

90].In addition, Lander & Kruglyak proposed a classification of results according to the level of significance obtained in a whole genome scan, being either suggestive (evidence expected one time at random in the genome scan), significant and highly significant linkage (5% and 0.1% p-value for the entire experiment).

Another related way of determining significance threshold is to simulate data. Instead of picking observations at random among the real data, a theoretical distribution (usually following a Gaussian distribution) is simulated and the threshold is determined as for permutation.

Even if robust, permutation and simulation requires a lot of computation, meaning a lot of time especially when many tests are performed.

5.3.4.2 Defining boundaries for the QTL position

In addition, due to the LD pattern, high test statistics values can be obtained on several centimorgans (cM) hence making difficult to ascertain the QTN position especially in LA. Therefore a confidence interval in which the probability of finding the QTL is above a certain threshold, is usually estimated [325]. In their influential paper, Lander & Botstein proposed to take the QTL confidence interval as every position for which the computed statistics was above the maximal statistics minus one LOD. This is a so-called "one lod drop-off" method. When working with LRT statistics, the 1 LOD drop-off becomes 4.6 LRT drop-off (as LR = 2ln(10)LOD) [325].

Under asymptotic conditions that are fulfilled for QTL with strong effects, the previous criterion provides confidence interval containing the QTL with 96.8%. However for QTL with weaker effect or for small populations the 1-LOD interval will have a probability of containing the QTL lying between 60% and 95%, while dense marker map also biases downward the 90% confidence interval [335]. Van Ooijen also stated that a 2-LOD interval be considered in a simulation study [399]. Another way to estimate the confidence interval is to perform bootstrapping as proposed by Visscher *et al.* [539]. During this process, bootstrap samples are created through sampling of N out of the N available observations with replacement, some observations being samples several times and other never. After n samplings, empirical confidence interval is estimated by taking the top and bottom 5th percentile (for a 90% interval) of the ordered estimates and the process is repeated a high number of times. Authors reported predicted confidence intervals close to the expected values [539]. However this approach is rather conservative [537] and also computationally demanding. In addition, a recent study by Manichaikul *et al.* demonstrated that the bootstrap confidence interval estimates vary widely according to the location of the

QTL relative to the genetic marker, *i.e.* at or close to genetic markers with respective probability to actually contain the QTL of 99% and 92.5%. This is due to the tendency of QTL position estimates to map precisely at a genetic marker [547].

5.3.5 Correction for the rest of the genome

Complex traits are under the control of many genes with various effects. While mapping a QTL, only a restricted fragment of the genome is considered thus omitting the effect of the rest of the genome on the given trait. This can lead to rejecting the null hypothesis too frequently, as demonstrated by Visscher & Haley (1996) [538].

Some methods have been proposed to consider the presence of another QTL segregating on the investigated linkage group. Jansen developed a two-step method by first selecting markers contributing the most to the trait of interest and subsequently performing an interval mapping with the selected markers as cofactors [250]. Jansen outlined the related issues of this method as being the high computational work required and the problem of selecting the right genetic model [250]. Subsequently Zeng proposed a similar method of composite interval mapping in which a classical interval mapping is performed with linked markers fitted to the model [576]. Subsequently Kao *et al.* extended the problem to multiple intervals containing putative QTL that are simultaneously considered in the model [259]. This model also considers epistasis effect between QTL but it is largely hampered by its computational requirements and also addresses a supplementary question about how to declare significance of the mapped QTL [259]. A full comparison of methods detecting multiple QTL have been published by Goffinet & Mangin in which they stated that methods starting with two QTL model are usually more powerful [189].

Another way to consider the genome of an individual as a whole is to perform a variance component analysis as suggested by Goldgar [190]. The phenotypic value z_i of an individual i is partitioned as follow [190, 325]:

$$z_i = \mu + A_i + A_i^* + e_i$$

where μ is the population mean, A_i is the proportion of variance contributed by the considered interval and A_i^* is the contribution outside the interval and e is the residual. Not only considering the genome as a whole, this method also offers the advantage of considering complex pedigrees as encountered in animal populations. However the major drawback is not being able

to map rare variants as the genetic variance is conditioned upon the QTL allelic frequencies [325].

Since many more markers are now available, the effect of the rest of the genome can be accounted for by estimating a genomic value by using all other SNPs than the tested one. This is the principle of the genomic selection described in more detail in section 5.6.2.

5.3.6 Is there any ideal method for QTL mapping ?

A QTL mapping experiment should answer a biological question that usually is "where are the genes controlling the trait of interest ?". Hence it is not only required to detect every QTL affecting the trait but also to precisely locate them.

In the case where few markers are available, *e.g.* microsatellite panels, the linkage analysis is the method of choice. Its precision is mostly conditioned by the number of observed meiosis provided that a good marker density is achieved (1 marker per 10 cM) [106]. In addition, this method is particularly robust, as the markers inherited by the progenies from their sires are well traced, which provides a good knowledge to assess the presence of a QTL segregating within the family. In a first attempt one should put the effort on increasing the number of observed families rather than having big families. This will allow for screening a larger amount of alleles in the population. After this first mapping, segregating families should be increased in size to produce more recombinant animals and to subsequently refine the position of the already detected QTL [105].

If many markers, *e.g.* DNA SNP chip, are available, LDA and LDLA that exploit existing LD is the method of choice. In addition,the LDLA methodology appears to be more robust than the LDA methods even if sometimes more computationally demanding. However this advantage also depends on the considered experimental design as demonstrated by Lee & Van der Werf (2004) [305]. Exploiting LD across a population requires an exhaustive list of haplotypes segregating in this population as well as a minimum of observations per haplotype so that corresponding effects can be correctly estimated [305]. This is main conclusion from the simulation study by Lee & Van der Werf who reported little advantage of LD mapping when there are few families with large size [305].

The power of a QTL mapping experiment can be increased by genotyping animals with extreme phenotypes only in a so-called "selective genotyping" design [292]. This method aims at targeting the animals responsible for most of the observed variance while reducing the number

of genotypings to be performed [292]. In that case, a correction should be applied on QTL effects estimated on correlated traits [65]. Another solution is the grand-daughter design suggested by Weller *et al.* that is using the genetic merit of progenies as a trait for performing QTL mapping [551]. This was implemented in dairy cattle.

5.4 Looking for selective sweep can help pinpointing QTL

5.4.1 Definition [213, 390]

A selective sweep is the result of an increased local selection of the genome due to the appearance and the rapid selection of a particular mutation. Such sweep usually greatly affects the surrounding genomic region especially through the "hitch-hiking phenomenon" as the neutral markers tightly linked to the region under selection will be passed through the generations. Therefore detection of a selective sweep provides knowledge about which genomic regions are under selection and can help pinpointing interesting QTL regions, provided that both regions co-localize [389, 60, 554, 397].

5.4.2 Different tests for detection [390, 389, 397]

Provided that most polymorphisms are selectively neutral [276], neutral theory constitutes a null model to be opposed to the specific occurrences of selection [389]. Various tests have been proposed that focus either on the allelic distribution and the levels of variability, or on the comparison of variability in different classes of mutations. Other additional tests also try to model selective sweeps especially by considering the LD pattern around selective sweep.

5.4.2.1 Within population tests

The first test to be developed was proposed by Lewontin & Krakauer in 1973 [312]. This test has subsequently been tuned over the past few years in parallel of the advances with genomics [390]. The Lewontin & Krakauer method considers several loci and aims at estimating the variance in allelic frequency of a given population. If the variance is too high, the neutral null model can be rejected.

In this lineage, the frequency spectrum that is the count of the number of mutations with a given frequency, can also be considered within population. Many of the developed tests focus on this frequency spectrum. Among these the Tajima's D test is one of the most famous : it

tests if the difference in variability between pairs of sequence is larger than what is expected on the standard neutral model. If so, a sweep is detected [499].

Above cited tests usually rely on strong assumptions on demographic properties. For instance, Tajima's test includes assumptions of constant population size with no population structure. One way to overcome this has been proposed by Kim & Stephan [275]. Considering the particular pattern of variability surrounding the sweep region, they integrated the distance of a particular site to an advantageous mutation in the computation of the frequency spectrum. In the end they could estimate the location and the strength of the sweep. Other methods take in consideration the LD increase occurring in selected regions, especially for incomplete sweep that have not reached fixation [448]. Sabeti *et al.* proposed to have a look at long-range haplotypes in human populations for which they estimate their respective ages by the decay of their associations to alleles at various distances from a locus of interest. This is measured by the extended haplotype homozygosity coefficient (EHH). Any higher EHH value is synonymous of an increase of a mutation frequency in the gene pool faster than expected under the neutral theory [448]. This test has been extended for cross population study (called XP-EHH) to detect selective sweeps occurring in a particular human population but remaining polymorphic in the human population as a whole [449]. These LD-based tests for detecting sweeps can be sensitive to assumptions regarding recombination rates.

5.4.2.2 Comparative data

While already-mentioned tests focus on ongoing selection and are particularly sensitive to demographic assumptions, other tests use species comparison and are hence free of such working hypothesis. Such comparison are used to identify older events of selection [397].

A first approach is observe the between-species divergence of homologous segments. This phylogenetic shadowing takes into account the evolutionary context.

A most straightforward test for demonstrating positive selection is to estimate the ratio between the number of non-synonymous substitutions per non-synonymous site (d_N) and the same ratio for synonymous sites (d_S). If the ratio $\frac{d_N}{d_S}$ is larger than one then non-synonymous mutations tend to be favored suggesting positive selection whereas a ratio below one speaks in favour of negative selection.

The MacDonald-Kreitman test counts the number of non-synonymous and synonymous mutations [347]. The ratio between of non-synonymous to synonymous fixations between species is

compared to the ratio of non-synonymous to synonymous polymorphisms within species. In the absence of selection, the ratios should be the same for both synonymous and non-synonymous. In case of a non-synonymous mutation occurring, selection will either reduce or increase the frequency of this mutation according to the respective detrimental or beneficial effect it can have, hence shifting the related ratio of non-synonymous mutations. Provided the sites that are affected by mutations are homogeneously spread over the genome they are similarly affected by changes in population size or drift. A similar test, named Hudson-Kreitman-Aguade, has been developped by Hudson *et al.* that contrasts polymorphism and divergence among multiple loci [241].

5.4.3 A few examples

Livestock have been selected to adapt to a broad range of environments. Hence some regions of the genome should have been/be under selection and their detection could help pinpointing gene of interest. The advances of genomics have been exploited in recent studies that searched SNP associated to high values of the Wright's F_{st} coefficient, that is a measure of genetic differentiation between populations [554].

In the frame of the sheep HapMap project, 2,819 sheep belonging to 74 breeds were genotyped with the Illumina 50K SNP Beadchip [274]. Not only investigating genetic diversity among breeds as well as their genetic relationships, a genome scan was carried out for F_{st} coefficient [274]. F_{st} coefficient was estimated for each breed related to other breeds. This analysis revealed 31 genomic regions summarizing the top 0.1% of the highest F_{st} scores. These regions contain genes related to coat pigmentation, skeletal morphology, body size, growth, and reproduction and the major signal was found near the relaxin/insulin-like family peptide receptor 2 (*RXFP2*) located on OAR10 and associated to the polled phenotype [274].

Interestingly no genes related to immune response or resistance to disease were found in the identified regions. This may be due to the lower selection pressure pathogens represent in sheep breeding as chemical therapeutics can overcome this pathogenic load. However Dayo *et al.* tried to look for selection sweep related to trypano-tolerance [110]. They genotyped 92 microsatellite markers for 509 cattle split into four West African trypano-tolerant taurine breeds and 10 trypano-susceptible breeds. Interestingly they found that BM4440 and DIK5250 had a significant reduced variability in the West African cattle populations (group defined according to the genetic structure of breeds) with trypano-tolerant status. Even if it is likely that the

underlying mutation is rather a locus linked to the markers than the marker itself, *CXCR4* lying at 500 Kbp away from BM4440 was proposed as a functional candidate.

Another study of the same type compared nine different fat- and thin-tail sheep breeds (Zel and Lori Bakhtiari breeds from Iran and other breeds from the sheep Hapmap dataset) to look for genomic regions under selection controlling fat deposition [373]. They found seven candidate regions with larger values in the Iranian set of breeds, three of which were found in the sheep Hapmap dataset as well and located on OAR5, 7 and X [373]. Additional investigation on the mean homozygosity of the three candidate regions revealed that homozygosity was increased on chromosome 5 and X for Lori Bakhtiari (fat tail) and at the candidate region on chromosome 7 for Zel (thin tail). Interestingly the region on OAR5 had already been related to fat thickness in cattle [373].

5.5 QTL mapped for resistance to GIN in sheep (table 5.4)

There have been a wide range of studies for mapping genes underlying resistance to GIN in sheep [38, 101, 108, 42, 203, 336, 128, 270, 342, 469].

Most of the studies have used microsatellites markers except a work published by Kemper *et al.* (2011), which was more a genomic selection approach [270]. QTL have been found on many chromosomes which is certainly linked to the variations between studies in terms of populations, nematode species and number of markers (see table 5.4). Still, OAR3 and OAR20 have been particularly common in QTL mapping studies for resistance to GIN and *IFN*γ and *MHC* are evident functional candidates identified on these two chromosomes respectively [127, 49].

The usually recorded trait for QTL detection are FEC as they provide a good knowledge about phenotypic resistance and are relatively easy and cheap to collect. Some studies have focused on more refined phenotypes like Davies *et al.* [108] who recorded IgA concentration as well as worm length and worm burden during a *T.circumcincta* infection or Dominik *et al.* who monitored changes in eosinophil number after *T. colubriformis* challenge [128]. In the same way an extensive phenotyping of controlled haemonchosis (worm burden, worm female length, IgG concentration in abomasal mucus or in plasma, abomasal pH and pepsinogen concentration) performed in BC animals with extreme resistant/susceptible phenotypes was reported by Moreno *et al.* (2006) [374].

Until the publication by Kemper *et al.*, QTL effects were found to be moderate. Interestingly

Table 5.4: QTL found for resistance to GIN in sheep [112]

OAR	Population	Strongyles	Trait	Candidate gene	References
1	Merinos, Romneys,Spanish Churra	Mixed, Tcol, Hcon	FEC, WB, IgA		[38, 122, 108, 203, 336]
2	Scottish Blackface,SardaxLacaune, Merinos, Romneys	Mixed, Hcon	FEC, WB		[374, 101, 336]
3	Scottish blackface,Merinos, Soay, SardexLacaune, BlackBellyxINRA 401, Texel	Mixed, Tcol, Hcon, Tcir	FEC, IgA	INFG	[94, 38, 459, 108, 374, 336]
4	SardaxLacaune	Mixed, Hcon	FEC		[374, 336]
5	BlackBellyxINRA401, Corriedale, Polwart	Mixed, Hcon	FEC, PCV	IL-3, IL-4, IL-5	[39, 374]
6	Merinos, SardaxLacaune, Spanish Churra	Mixed, Tcol, Hcon	FEC	GR01, KIT, IF1	[38, 374, 203, 336]
7	Black bellyxINRA401, merinos	Hcon	FEC		[374, 336]
8	SardaxLacaune, Merinos, Romneys	Mixed, Hcon	FEC, WB		[374, 203, 336]
9	Merinos	Hcon	FEC		[336]
10	SardaxLacaune, Spanish Churra, Merinos	Mixed, Hcon	FEC		[374, 203, 336]
11	Merinos, Romneys	Mixed, Tcol, Hcon	FEC, WB		[38, 101, 336]
12	Merinos, SardaxLacaune, BlackbellyxINRA401	Mixed, Tcol, Hcon	FEC		[38, 101, 336]
13	SardaxLacaune, BlackbellyxINRA401	Mixed, Hcon	FEC, PCV		[374]
14	Scottish Blackface, SardaxLacaune, Spanish Churra	Mixed	FEC		[108, 374, 203]
15	Merinos	Hcon	FEC		[336]
16	SardaxLacaune, Merinos	Mixed, Hcon	FEC		[374, 336]
18	Merinos	Tcol, Hcon	FEC	IgE	[91, 336]
19	SardaxLacaune	Mixed	FEC		[374]
20	Scottish Blackface,Rhoenschaf,Polish Heath, Soay, Merinos, Suffolk	Mixed, Hcon, Tcir	FEC, PCV, IgA, eosinophils	MHC	[467, 76, 408, 252, 458, 108, 336]
21	SardaxLacaune, Merinos	Mixed, Hcon	FEC		[374, 336]
22	Merinos	Hcon	FEC		[336]
23	Black-bellyxINRA401, Romneys	Mixed, Hcon	FEC, IgG, IgE		[374, 101]
24	Merinos	Hcon	FEC		[336]
25	Merinos	Hcon	FEC		[336]
26	Merinos	Hcon	FEC		[336]
x	Merinos	Hcon	FEC		[336]

Key: Hcon: *H. contortus*; Tcol: *T. colubriformis*; Tcir: *T. circumcincta*; WB: worm burden

a major gene effect for resistance to *H.contortus* have been reported but has never been confirmed using genetic markers [361] in [520].The use of genomic selection methods to detect genetic markers explaining most part of the variance showed that QTL exhibited low effects [270].

Interestingly, two QTL on OAR3 and OAR14 (out of the three significant QTL) mapped by Davies *et al.* in a Scottish Blackface population [108] were found by Matika *et al.* in two other breeds (Texel and Suffolk), reinforcing the actual presence of a QTL on these chromosomes [342].

5.6 Integration of molecular markers for sheep breeding

Classical selection schemes mostly rely on the use of phenotypes and pedigree to predict the genetic merit of each candidate. However targeting the genes of interest with genetic markers could increase the genetic gain, especially for traits with low heritability, for traits being under the control of one or more QTL or if the candidates cannot express the trait of interest [479, 290]. The use of molecular information can aimed either at introgressing a gene from one population to another or to take advantage of QTL knowledge to better estimate the genetic value of an individual. This marker-assisted selection (MAS) can relied on three different genetic information [118]:

- LD-MAS that exploits the linkage disequilibrium between the QTL and the markers across population
- LE-MAS, for which marker genotype and QTL are not tightly correlated so that inheritance must be traced within the pedigree structure
- Gene-MAS

In addition to these approaches, it is also possible to estimate the genetic merit of an individual by considering a great number of loci. This has been made possible thanks to genomic data and have given birth to the genomic selection approach (GS). MAS and GS are reviewed under a sheep perspective in the frame of resistance to GIN.

5.6.1 Marker-Assisted selection [118, 124]

The basic theory of MAS has been proposed by Neimann *et al.* [386] and Smith [476] and the general implementation into breeding schemes has been discussed by Lande & Thompson [290].

MAS can be implemented in three different ways, either through a tandem selection that first targets candidates based on their genotype and is followed by selection on genetic merit, or on the basis of an index integrating both marker information and genetic value or a pre-selection based on marker at a young age followed by a selection based on phenotype or genetic merit later in life [118].

The inclusion of genetic markers into BLUP breeding values has been derived by Fernando & Grossman [156] and simulations by Meuwissen & Goddard have estimated a genetic gain of 8 to 38% [363].

LE-MAS require more logistic for routine evaluation. Indeed, QTL ascertainment is done on a within-family basis so that phenotyping and genotyping of selection candidates and their relatives is required. This type of breeding scheme has been successfully implemented in the dairy cattle genetic estimation routine in France until 2008 [58, 202] where it became a mixture of LD-MAS and genomic selection (A. Legarra, personal communication).

Both LE- and LD-MAS will provide lower extra genetic gain than direct markers due to losses in accuracy. LE-MAS is hampered by the within family frame that limits the number of available observations [536] while LD-MAS is affected by incomplete LD between QTL and markers and the high number of effects to be estimated. To this regard, Hayes *et al.* demonstrated that the use of haplotypes for MAS purpose could increase accuracy of selection if the good number of markers could be chosen as the more markers included in the haplotype, the more haplotypes and the more effects to be estimated which in turn reduces the accuracy [218].

Such selection approach has not warranted sheep industry testing schemes, as benefits would not have overcome related costs that are DNA collection, genotyping and analysis [124]. The release of an ovine genomic SNP chip has considerably modified this position as developed in the next paragraph.

5.6.2 Genomic selection: concept and prospects for breeding sheep resistant to GIN

One of the reference work about genomic selection has been proposed by Meuwissen, Hayes and Goddard in 2001 [362]. It aims at taking advantage of dense marker maps to directly estimate the breeding value from the DNA information: gEBV (genomic Estimated Breeding Value) [188]. Briefly, a reference population is constituted and recorded for the trait(s) of interest and genotyped for a dense set of markers. From these phenotypes and genotypes, the effect

of each genotyped marker can be estimated and combined into a prediction equation to finally estimate the genomic merit of each individual. This prediction equation can then be applied to any genotyped individual with no need of a phenotype [188]. This is of particular interest as the relationship between the marker and the QTL will hold across population and that almost every QTL should be in LD with at least one of the genotyped marker [362, 188, 187]. Genomic selection is thus an extension of the LD-MAS at a genome wide scale [364].

In classical genetic selection, genetic merit of selection candidates (EBV) are evaluated through the phenotypes of related individuals and the pedigree information using BLUP (Best Linear Unbiased Predictor) estimates. Therefore, selection is a long process. Genomic selection revolutions this scheme as prediction of gEBV only relies on genetic markers, after the effect of each genetic marker has been estimated in a reference population [188]. Therefore genomic selection can spur genetic gain through increase of the selection pressure (especially for traits with low heritability) or decrease the generation interval (especially in cattle) and can help selecting for traits that are difficult to measure in routine [290, 188, 379]. However, this also addresses statistical challenges as the number of estimated effects is greater than the number of phenotypes. In their foundation paper, Meuwissen *et al.* compared three methods being a three-step least square estimation procedure and two Bayesian methods to the classical BLUP estimation based on pedigree information [362]. The least square procedure consisted in estimating the effects of haplotypes along the genome, selecting the apparent QTL and estimate the effect of each of the encountered haplotype at the QTL position before summing all values. Bayesian approach relies on the Bayes theorem that can be written as follows:

$$P(x|y) \propto P(y|x)P(x)$$

where $P(x|y)$ is the posterior probability (knowledge after the experiment is run), $P(y|x)$ is the likelihood and *P(x)* is the prior probability (before the experiment is run). In the genomic selection developed by Meuwissen *et al.* [362], not only the phenotypes are modeled but also the variances of each considered haplotypes. Related to the above equation, the posterior probability is the variance distribution knowing observed genotypes. The prior distribution of variances was taken as an inverted chi square that offered the advantage to result in a known inverted chi square distribution for the posterior probability. Genotypes effects were sampled from the normal distribution. This method, called BayesA started the Bayesian alphabet of genomic

selection methods [362, 178, 205]. A derived method presented in the same paper, BayesB was also implemented to take into account the fact that many SNPs have no effect on the considered trait so that some haplotypes will not contribute to the variance :

$\sigma^2 = 0$, with probability π

$\sigma^2 \sim chi^{-2}(\nu, S)$, with probability (1 - π)

From their results, BayesB fitted the best reality with an accuracy of 85% whereas least square methodology gave the worst results.

Extension of the Bayesian alphabet made by Habier *et al.* aimed at changing the prior distribution of markers variances by considering the π probability as an unknown parameter, either common to every loci (BayesC) or with a varying scaling parameter of the inverted χ^2 distribution (BayesD) [178, 205]. The authors concluded that none of these methods outperformed every other for each of the considered trait and that BayesA should be a good choice. However the dynamic research surrounding the development of genomic selection methods, either ridge regression [553], Bayesian procedures [362, 179, 570, 113], machine learning [321], partial least square (PLS) or sparse PLS [93] underlines well the need for precise and computationally efficient modeling of the genes effects to get estimates with best accuracy [364].

Genomic selection has proved to be useful in case of low heritability traits. Empirical results have proven the superiority of genomic selection over BLUP estimates, even if a simulation study by Muir (2007) have found opposite findings [379]. Muir thus proposed that genetic markers cannot map every QTL affecting the trait of interest so that gEBV computation should be done by considering both type of information [379, 217].

The constitution of the reference population is the major limiting factor before implementation as more than 1000 animals are required to provide good estimates [188]. In addition, this reference population needs to be continually updated. For sheep, Daetwyler *et al.* reported accuracies ranging from 0.15 to 0.79 for wool traits and from -0.07 to 0.57 for meat traits while Kemper *et al.* published the first estimates of gEBV for FEC [270]. However both studies concluded that reference population should be increased to both allow across breed estimation. In addition Kemper *et al.* estimated that QTL affecting FEC exhibited low effects so that marker density should be increased to get closer [270]. In France, computation of gEBV for resistance to GIN is mostly of interest in meat breeds or in dairy breed from Pays Basque in southern France. However it will be difficult for these breeds to constitute a big enough reference set for valuable accuracy in genomic estimates.

5.7 Bypassing QTL mapping with Next-Generation Sequencing [176]

The major issue with QTL mapping is to refine the rather large confidence interval to the looked for QTN [176, 442, 552].

5.7.1 Whole Genome Resequencing

Whole-genome sequencing data should help bypassing the QTL mapping step [176]. A recent publication by Larkin *et al.* used both re-sequencing data from two related bulls, *i.e.* father and son, contributing to 7% of the actual dairy cattle genetic diversity [294]. Using 50K SNP chip data from more than thousand progenies, the authors could identify 49 regions being under selection. Finally thanks to produced sequencing data, a limited number of functional candidates affecting milk production, reproduction traits and resistance to disease were identified. Provided the sequencing costs keep getting lower and lower such an approach might be routinely adopted in the next years. The first draft of the ovine sheep genome has now been available for a couple of years [18] and re-sequencing of individuals selected for a particular trait on the basis of their SNP genotypes will be soon performed either in Australia or in Europe.

5.7.2 eQTL

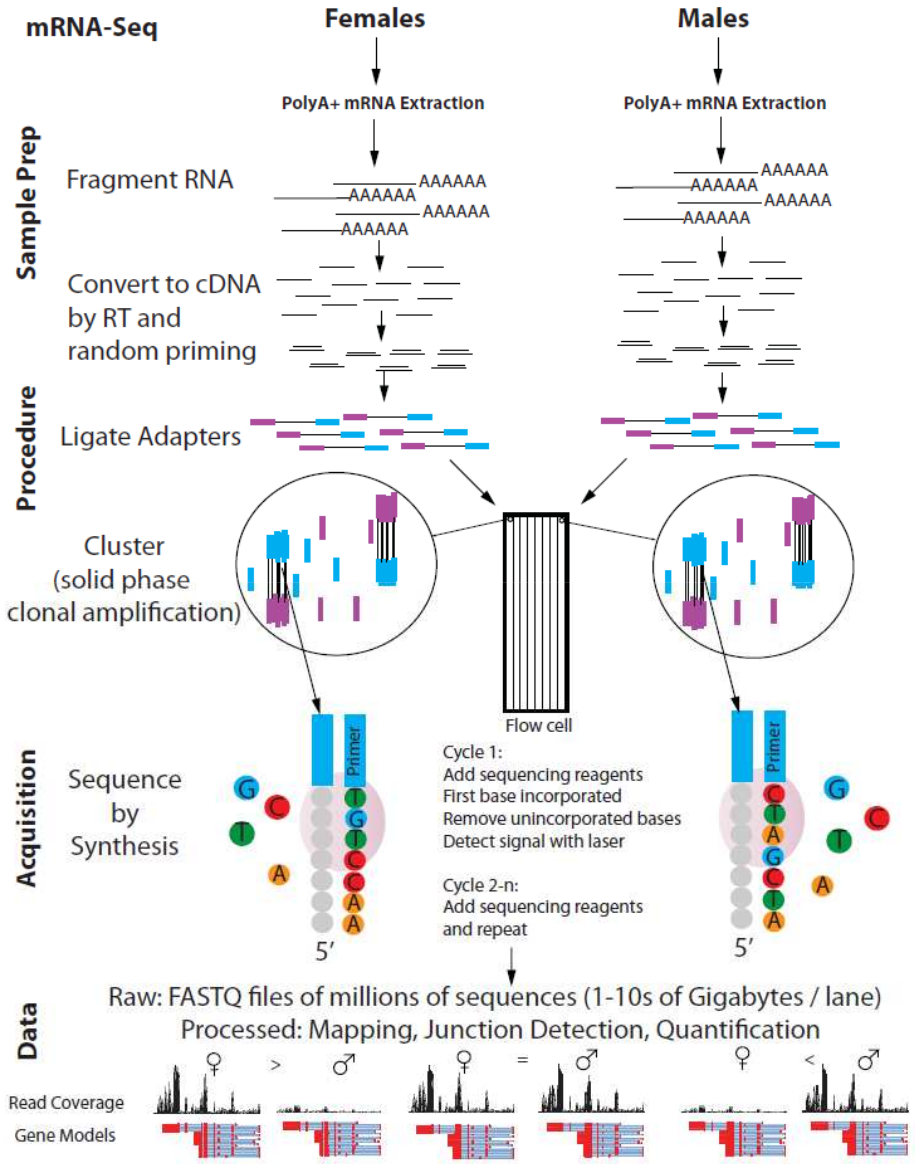
One way to overcome the problem of translating positional candidate into a functional gene, is to lead QTL detection and expression studies in parallel. Jansen & Nap [251] proposed to detect expression-QTL (eQTL) by combining genome-wide linkage analysis with expression studies. This approach would then allow a more global view on networks of gene implicated in disease resistance.

However, the high level of false positive detection and the lack of power of such eQTL design in livestock species are major issues. Moreover, such experiment requires hundreds of expression levels to be monitored and thus remains extremely costly [111]. So far as we know, only one recent study focused on bringing together expression data and quantitative genetics [255].

5.7.3 RNAseq

The problematic associated to the eQTL could be resolved by directly getting knowledge of what is expressed and where it lies. Thanks to the rapid evolution of next generation sequencing

Figure 5.3: Schematic representation of the RNAseq data production workflow (reproduced from Malone & Oliver (2011) [334])



methods, it has been possible to simultaneously access both information [548]. In principle, a RNA collection is converted to cDNA fragments with addition of adaptors at one or both ends of the segment for subsequent high-throughput sequencing (see figure 5.3). Produced reads are then aligned to a reference genome or transcriptome. Therefore, RNA-seq has the major advantage of providing exhaustive quantification of the available RNA with no need of a prior known genomic sequence as in microarray. In addition, identified DNA sequences can be mapped on the genome and the technique can identify either high or low range of expression levels (from 9,000 to 5 fold change in gene expression) [548]. Still some genes are impossible to detect with RNAseq approach [334]. One major limitation consists in the library creation as large RNA must be cut into smaller segments (200-500 bp) to be compatible with most deep-sequencing technologies, hence resulting in technical biases. In addition, the sequencing step results in producing a huge amount of data [334] that represents a bioinformatical challenge as sequence reads are expected to map several positions on the genome. In addition alternative splicing can make it difficult to map reads spanning splice junctions but recent developments enabled longer reads length hence providing more information [404]. In addition to the high costs the technology represents in comparison to microarray (100 versus 1,000 $ per sample respectively [334]), it should take time before RNAseq becomes the method of choice [334]. So far as we know, two papers using RNA-seq technology on sheep tissue have been published [253, 405] and one experiment is currently undergoing at the Toulouse INRA center (A. Bonnet, personal communication).

5.7.4 First paper: Application of genomic tools for breeding small ruminants

série la sélection génomique

apports de la génomique en élevage des petits ruminants

Guillaume Sallé[1,2]
Carole Moreno[2]

1. INRA, UMR1225, Interactions Hôtes-Agents pathogènes, 31076 Toulouse

2. INRA, UR631, Station d'Amélioration Génétique des Animaux, 31320 Castanet-Tolosan

Les nouveaux outils moléculaires permettent désormais d'évaluer la valeur génétique des ovins, et bientôt des caprins, à partir de leur seul génotype. Cette avancée majeure va révolutionner la sélection des petits ruminants, et pourrait favoriser l'intégration de la résistance génétique aux maladies dans les programmes de sélection. Cet article fait suite aux deux articles publiés dans *LE NOUVEAU PRATICIEN VÉTÉRINAIRE élevages et santé* n°16 sur les principes de la sélection génomique chez les bovins laitiers, il aborde la même thématique chez les petits ruminants.

Objectifs pédagogiques

- Connaître les méthodes de sélection.
- Comprendre les apports de la génomique.
- Comprendre les opportunités offertes en santé des petits ruminants.

Essentiel

- La génomique permet d'identifier précisément quels segments chromosomiques sont portés par un individu, et d'estimer l'effet de ces segments sur un caractère donné.
- La génomique accélère le progrès génétique et ouvre la voie à la sélection de nouveaux caractères d'intérêt.

NOTE

* ISGC : *International sheep genome consortium* http://www.sheephapmap.org/

RUMINANTS

Crédit Formation Continue : 0,05 CFC par article

Il est probable que la médecine vétérinaire doive restreindre l'usage des traitements chimiques dans les années à venir. Les attentes sociétales de produits "sains" et "respectueux de l'environnement" se traduisent par un encadrement plus strict de la fabrication et de l'usage des médicaments vétérinaires [8]. De plus, des souches de nématodes multi-résistantes aux anthelminthiques ont été recensées sur les cinq continents [12]. Menace endémique des élevages de petits ruminants en Amérique latine ou en Afrique du sud [14], ces souches multi-résistantes apparaissent en Europe et ont conduit à la disparition d'un élevage de 90 brebis et 130 agneaux en 2005 en Ecosse [21], puis d'un second en 2007 en Angleterre, où 60 agneaux ont montré des signes de résistance aux traitements [5]. Ces cas de multi-résistances sont d'autant plus critiques que très peu de nouvelles molécules anthelminthiques spécifiques des petits ruminants ont été mises sur le marché, ces espèces ne présentant pas un intérêt suffisant pour les laboratoires pharmaceutiques [3]. Cette situation est représentative de l'insuffisance de la pharmacopée vétérinaire pour les petits ruminants [8].

1 Les Martinik Black-Belly ont une résistance avérée à l'infestation par les strongles gastro-intestinaux, dont *Haemonchus contortus* (photo P. Jacquiet).

- **La génétique représente une alternative séduisante pour la maîtrise sanitaire dans les cheptels de petits ruminants** comme l'illustre la lutte contre la tremblante, maladie pour laquelle la seule mesure de contrôle disponible est l'abattage préventif des troupeaux atteints [11].
- **Dans le même temps, des avancées majeures ont été réalisées en génétique, permettant désormais d'explorer le génome à l'échelle de l'ADN. La génétique est ainsi entrée dans l'ère de la génomique.** Cela permet d'accélérer le progrès génétique en prédisant précocement la valeur génétique des individus à partir de leur ADN, et facilite l'identification des gènes affectant les caractères.
- Cet article présente les apports de la génomique en élevage des petits ruminants, et envisage les opportunités liées à l'intégration de la génétique pour maîtriser la santé du cheptel.

La génomique peut être utilisée à des fins de maîtrise sanitaire :

1. à l'échelle d'une race, en intégrant la résistance à certaines maladies dans les évaluations génétiques des mâles ;
2. à l'échelle du troupeau, avec une gestion fine de la génétique par l'éleveur.

LA RÉVOLUTION GÉNOMIQUE : LES NOUVEAUX OUTILS DISPONIBLES

Le développement d'une puce génomique

- Si les méthodes de domestication se sont complexifiées, le principe général de la

sélection, appuyée sur les performances propres des animaux pour sélectionner les meilleurs reproducteurs, est resté le même.

• *A contrario*, l'évolution des connaissances sur l'ADN, support des gènes, a été vertigineuse lors des 50 dernières années, même si les applications en sélection ont été quasi nulles jusqu'en 2009.

• **Depuis 2009, le travail effectué par le consortium international de génomique ovine* a permis le développement d'une puce génomique comprenant 54 241 marqueurs de type SNP** *(single nucleotide polymorphism)* ***(encadré)***.

La construction d'une puce comparable chez les caprins est en cours de réalisation.

Un double intérêt

• La révolution apportée par ces puces SNP permet d'accéder à un nombre de polymorphismes de l'ADN suffisant pour estimer l'effet de l'ensemble des gènes de l'animal**. Ceci offre un intérêt en sélection.

• De plus, en offrant une information en tout point du génome, ces puces ouvrent la voie à une localisation fine et à une identification des gènes impliqués dans le déterminisme des caractères d'intérêt, comme la résistance aux maladies.

LA GÉNOMIQUE ACCÉLÈRE LE PROGRÈS GÉNÉTIQUE

Les débuts d'application sur les mammites et sur les infections par les strongles

• **La sélection génétique classique d'animaux résistants aux maladies est une réalité chez les petits ruminants.**

Les travaux de recherche ont montré que la susceptibilité à certaines maladies était d'origine génétique, comme illustrée par des différences de résistance entre les races ***(photo 1)***.

• Chez la chèvre et le mouton, deux des principales affections, les mammites et l'infestation par les strongles gastro-intestinaux, sont étudiées d'un point de vue génétique [7] :

- on estime que 15 p. cent de la variation observée entre individus dans la résistance aux mammites est d'origine génétique [2] ;
- et en moyenne, la génétique explique un tiers de la variation observée entre individus dans la résistance aux strongles [4], c'est-à-dire autant que pour la quantité de lait produite. Cette variation est suffisante pour être exploitée en sélection.

• À l'heure actuelle dans la région de Roquefort, la sélection classique des béliers comporte déjà la résistance aux mammites prédite par les scores de comptage des cellules somatiques du lait [20]. Des études de faisabilité de la mise en place d'une sélection pour la résistance aux strongles gastro-intestinaux sont en cours sur des ovins laitiers au pays basque, et des ovins allaitants en Auvergne [13], en procédant à la mise à l'épreuve des candidats à la sélection.

• La sélection génomique au sens strict (*cf.* l'article *"La sélection génomique et son développement chez les bovins laitiers"****) repose sur l'association entre des marqueurs génétiques et une performance. Au sein d'une population dite "de référence", des animaux sont génotypés et mesurés pour des caractères d'intérêt, permettant ainsi de prédire l'effet statistique de chaque marqueur génétique ***(figure)***.

Les effets des marqueurs sont combinés dans les équations de prédiction, qui donneront la valeur génétique des individus à sélectionner, à partir de leur génotype, et sans mesurer leur performance [10].

— Encadré - SNP et puce génomique —

• Les marqueurs génétiques sont des fragments d'ADN, polymorphes à l'échelle d'une population *(tableau)*.

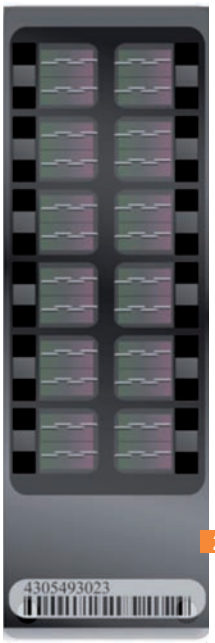
• **Parmi ces marqueurs, les SNP** *(single nucleotide polymorphism)* **sont des changements ponctuels de l'ADN :** la différence entre individus se situent sur une paire de base.

On trouve des milliers de ces SNP sur chacun des 26 chromosomes du mouton : le développement de SNP est en cours pour les caprins.

• **Le développement récent de puce à SNP permet de déterminer rapidement le génotype d'un individu *(photo 2)*.**

• Chez les ovins, cette puce comprend *in fine* 54 241 marqueurs.

- Pour connaître les différentes versions allèliques des 54 241 SNP portés par un animal, on extrait son ADN et on l'hybride sur la puce.
- Un automate permet ensuite d'accéder à la version de chaque SNP porté par l'animal.



2 Puce à SNP 50K ovine* (photo Illumina).

Tableau - Comparaison des outils de génétique moléculaire : Microsatellite *versus* SNP *(single nucleotide polymorphism)*

	Panel Microsatellite	Puce SNP
• Nombre de marqueurs	- 150 à 200	- 50 000 à 60 000
• Nombre d'allèles polymorphes	- 2 à 20	- 2
• Nombre de marqueur par cM	- 0,5	- 30
• Localisation	- Entre les gènes	- Dans et entre les gènes
• Prix	- 200 à 300 €	- 100 à 150 €*

* Ces prix sont appliqués lors de commandes groupées au niveau international, organisées une fois par an.

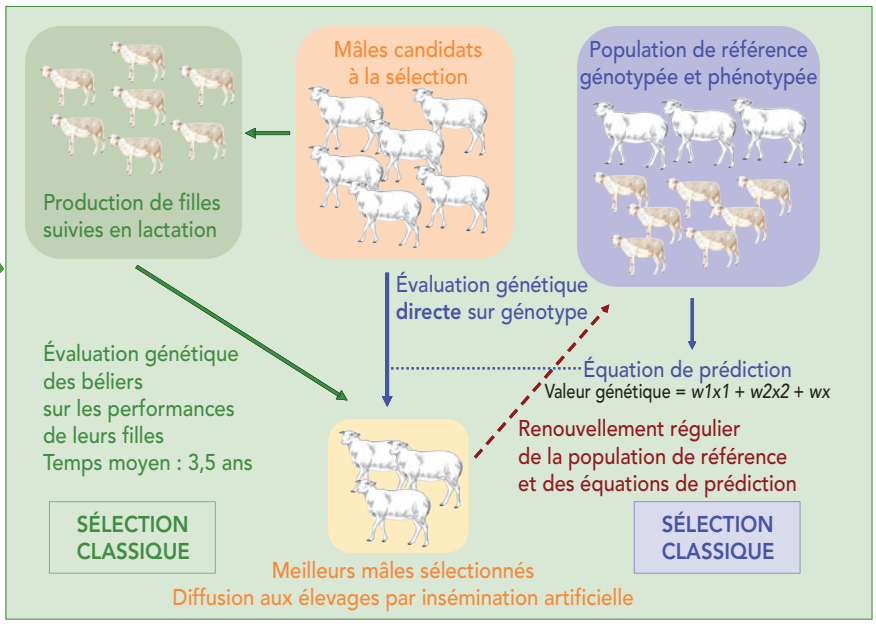
NOTES

* Commercialisée par Illumina http://www.illumina.com

cf. Les articles :

- ** *"Principes de l'évaluation génomique chez les bovins laitiers"*,
*** *et"La sélection génomique et son développement chez les bovins laitiers"*,
par D. Boichard, F. Guillaume, A. Baur, et coll., *LE NOUVEAU PRATICIEN VÉTÉRINAIRE élevages et santé* 2010;4(16):330-4.

Figure - Comparaison entre sélection classique et sélection génomique en race ovine laitière



- La sélection génomique permet une indexation directe des béliers, à partir de leur génotype et d'équations de prédictions préétablies.
- Ceci permet un gain de temps considérable par rapport au schéma classique qui nécessite un temps moyen de 3,5 ans.
- De plus, les candidats n'ont plus besoin d'être mesuré pour le caractère d'intérêt ; ceci facilite la sélection pour des caractères de résistance aux maladies qui pourraient altérer les performances des mâles reproducteurs.

Essentiel

- La génomique améliore l'efficacité de la sélection et facilite l'intégration de caractères difficiles à sélectionner (coûteux et/ou délicats à mesurer et/ou peu héritables).
- La sélection génomique au sens strict est très efficace dans les schémas laitiers, en race Lacaune notamment.
- La sélection pour la résistance aux maladies peut bénéficier des applications génomiques, si l'impact de cette sélection est mieux connu.
- La valorisation des informations sanitaires de terrain propres à chaque individu pourraient favoriser la sélection d'animaux résistants, et la découverte de gènes responsables de maladies.

RUMINANTS

Avantage majeur de la sélection génomique : des valeurs génétiques fiables dès la naissance ...

● **L'obtention de valeurs génétiques fiables dès la naissance des animaux est un des avantages majeurs de la sélection génomique, comparée à la sélection génétique classique.** L'intervalle de génération peut ainsi être réduit dans les schémas où la sélection des mâles repose sur des aptitudes mesurées sur leurs filles [10], comme par exemple les paramètres de lactation en élevage laitier ou les aptitudes maternelles pour les races bouchères. En outre, il est possible de réduire le nombre d'individus qui entrent en station de testage en déterminant à l'avance leur valeur génomique globale, ce qui se traduit par un gain économique non négligeable.

● **De manière concrète, la sélection génomique permet d'acquérir une valeur génétique dès la naissance de l'animal équivalente à celle obtenue en évaluation classique s'il avait produit 30 à 150 filles testées.** Une étude de faisabilité, qui comprend la quantification des bénéfices économiques apportés par la sélection génomique en race Lacaune, est en cours d'étude (les résultats seront disponibles en 2013).

● **La sélection génomique affranchit la sélection des individus de la mesure de performance, celle-ci n'est plus mesurée que dans la population de référence.** Il est donc possible d'intégrer dans les schémas de sélection de nouveaux caractères délaissés jusque là, en raison de leurs coûts de mesure élevés et/ou de leur impact sur la santé non négligeables. C'est ainsi que la sélection pour la résistance aux maladies bénéficiera probablement de ces avancées.

... mais une limite : la taille de la population de référence

● Toutefois, la sélection génomique se heurte à une limite, la taille de la population de référence : on estime qu'au moins 1000 individus sont nécessaires pour obtenir une valeur génétique estimée de précision suffisante [10]. Cette taille de population peut être atteinte en race laitière, Lacaune notamment. Au contraire, les effectifs des races allaitantes sont trop réduits pour constituer une population de référence dans chaque race. On pourrait envisager la création d'une population de référence multiraciale, mais les premiers résultats obtenus en génétique bovine semblent indiquer qu'un tel schéma n'offre pas, à l'heure actuelle, des précisions d'estimation suffisante.

● De plus, il est nécessaire de renouveler régulièrement la population de référence. En effet, les équations de prédiction estimées à un temps *t* dans une population de référence sont susceptibles d'être moins précises au bout de plusieurs générations sans phénotypage.

La sélection directe de gènes

● La sélection génomique envisage le génome d'un animal dans sa globalité sans connaître les gènes ni leurs modes d'action. Il est cependant possible de ne considérer qu'un nombre restreint de gènes connus.

● **C'est, par exemple, sur cette approche que repose la sélection pour l'éradication de la tremblante, dans sa forme classique, en France.** La susceptibilité à cette maladie est conditionnée par le génotype au gène PrP, qui code pour la protéine prion : l'allèle VRQ confère la sensibilité, alors qu'un génotype ARR/ARR offre la plus grande résistance à la tremblante classique, sans que l'on sache s'il s'agit d'un allongement de la durée d'incubation au delà de la limite de vie des animaux ou d'une résistance réelle [4, 9].
Le programme français d'éradication de la tremblante a permis la diffusion du génotype favorable et une diminution de la fréquence de l'allèle VRQ dans les noyaux de sélection [22]. Des études visant à déterminer le lien entre la sélection réalisée et l'incidence de la tremblante classique sont en cours.

● Pour le parasitisme ou les mammites, très peu de gènes sont formellement identifiés comme déterminant la résistance. Néanmoins, l'exploitation des nouveaux outils génomiques a déjà permis d'affiner de manière importante (quelques mégabases de l'ADN) les régions du génome impliquées. Une fois les gènes découverts et validés, leur utilisation est alors transposable à toutes les races.

LES APPLICATIONS POUR LA GESTION SANITAIRE DU TROUPEAU

● **L'utilisation de la génétique pour améliorer la résistance globale d'une race est un moyen de contrôle à long terme,** qui sera d'autant plus bénéfique que la prévalence de la maladie est élevée et que les traitements curatifs et/ou préventifs sont peu efficaces ou difficilement acceptables (abattage sélectif).

● **Toutefois, il est nécessaire de s'assurer qu'aucune contrepartie défavorable n'apparaisse, comme par exemple une chute de performances.** En effet, les animaux sélectionnés pour un niveau de productivité élevé sont certainement ceux dont le métabolisme privilégie le plus les caractères de production [6]. On peut alors se demander si des animaux sélectionnés pour une meilleure résistance aux maladies ne risquent pas de détourner l'allocation protéique vers la production de facteurs immunitaires au détriment d'autres fonctions physiologiques. Ceci pourrait expliquer les différences de conformation observées entre des races résistantes aux parasites et des races sensibles *(photo 3)*.

● **La sélection pour la résistance à une maladie ne doit pas non plus être synonyme de sensibilité à un autre pathogène.**

3 Les animaux de race Romane présentent une susceptibilité accrue en primo-infestation.

Cependant, en se basant sur trois lignées divergentes pour la réponse immunitaire (réponse en anticorps, réponse cellulaire et activation des macrophages), Pinard-van der Laan (2002) a montré que la sélection pour un seul de ces critères n'influe pas sur les autres [18]. Antérieurement, ce résultat avait été démontré chez la souris [16].

● **Par ailleurs, il est indispensable de s'assurer que les pathogènes ciblés ne contournent pas la résistance génétique qui leur est opposée.** Peu d'études ont été menées jusqu'à présent chez le mouton pour étudier ce phénomène. Toutefois, Kemper et coll. (2007) ont montré que les strongles ne semblent pas s'adapter à un hôte sélectionné pour leur être résistant [15].

● Cette non adaptation est observée chez les plantes et semble s'expliquer par la multiplicité des gènes en cause dans la résistance [19]. En effet, le facteur clé dans l'adaptation du pathogène à son hôte est l'hétérogénéité génétique de ce dernier.

- **Une résistance conditionnée par un seul gène à effet fort, entraîne une pression de sélection importante du pathogène qui conduit à la création d'une population composée des seuls variants susceptibles de contourner cette résistance.**

- **À l'inverse, une résistance sous le contrôle de plusieurs gènes requiert chez le pathogène un nombre élevé de mécanismes d'adaptation.**

● Dans l'hypothèse où de telles adaptations seraient développées, les coûts métaboliques associés pour le pathogène pourraient devenir un handicap majeur à la diffusion de ses gènes en milieu naturel [15].

Remerciements

Les auteurs remercient :
- R. Rupp et D. François pour la relecture de cet article et leurs commentaires constructifs, ainsi que S.C. Bishop pour ses avis éclairés ;
- les lecteurs arbitres du *NOUVEAU PRATICIEN VÉTÉRINAIRE élevages et santé*, pour l'ensemble de leurs remarques constructives ;
- C. Sidani pour les renseignements fournis sur le programme français d'éradication de la tremblante, ainsi que l'équipe "Génétique et Génomique des Petits Ruminants" de l'UR631.

Références

1. Collectif ANSES. Avis de l'Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail relatif à la politique de sélection génétique des ovins à long terme pour la résistance aux EST. Maisons-Alfort: Anses, 2010:1-18. .
2. Bergonier D, de Cremoux R, Rupp R, coll. Mastitis of dairy small ruminants. Vet Res 2003;34:689-716.
3. Besier B. New anthelmintics for livestock: the time is right. Trends Parasitol 2007;23:21-4.
4. Bishop SC, Morris CA. Genetics of disease resistance in sheep and goats. Small ruminant research 2007;70:48-59.
5. Blake N, Coles G. Flock cull due to anthelmintic-resistant nematodes. Vet Rec 2007; 161:36.
6. Coop RL, Kyriazakis I. Influence of host nutrition on the development and consequences of nematode parasitism in ruminants. Trends Parasitol 2001;17:325-30.

Suite p. 18

RUMINANTS

7. Davies G, Genini S, Bishop SC, Giuffra E. An assessment of opportunities to dissect host genetic variation in resistance to infectious diseases in livestock. Animal 2009;3:415-36.
8. Dehaumont P, Moulin G. Evolution du marché des médicaments vétérinaires et de leur encadrement réglementaire : conséquence sur leur disponibilité. Bull Acad Vét France 2005;158:125-36.
9. Fediaevsky A, Ducrot C, Calavas D. Surveillance épidémiologique des tremblantes ovines classique et atypique : enjeux et résultats. Le Nouveau Praticien Vét élevages et santé 2010;4(16):321-6.
10. Goddard ME, Hayes BJ. Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nature Reviews Genetics 2009;10:381-91.
11. Hunter N. Scrapie: uncertainties, biology and molecular approaches. Biochim Biophys Acta 2007;1772:619-28.
12. Jackson F, Coop RL. The development of anthelmintic resistance in sheep nematodes. Parasitology, 2000;120 Suppl:S95-107.
13. Jacquiet P, Barillet F, Bouix J, coll. La résistance génétique des ovins aux strongles gastro-intestinaux. Bull Acad Vét France 2009;162:39-46.
14. Kaplan RM. Drug resistance in nematodes of veterinary importance: a status report. Trends Parasitol 2004;20:477-81.
15. Kemper KE, Elwin RL, Bishop SC, coll. *Haemonchus contortus* and *Trichostrongylus colubriformis* did not adapt to long-term exposure to sheep that were genetically resistant or susceptible to nematode infections. Int J Parasitol, 2009;39(5):607-14.
16. Mouton D, Bouthillier Y, Mevel JC, Biozzi G. Genetic selection for antibody responsiveness in mice: further evidence for inverse modification of macrophage catabolic activity without alteration of the expression of T-cell-mediated immunity. Ann Immunol (Paris) 1984;135D:173-86.
17. Osteras O, Solbu H, Refsdal AO, coll. Results and evaluation of thirty years of health recordings in the Norwegian dairy cattle population. J Dairy Sci 2007;90:4483-97.
18. Pinard-van der Laan MH. Immune modulation: the genetic approach. Vet Immunol Immunopathol 2002;87:199-205.
19. Poland JA, Balint-Kurti PJ, Wisser RJ, coll. Shades of gray: the world of quantitative disease resistance. Trends Plant Sci 2009;14:21-9.
20. Rupp R, Lagriffoul G, Astruc JM, Barillet F. Genetic parameters for milk somatic cell scores and relationships with production traits in French Lacaune dairy sheep. J Dairy Sci : 2003;86:1476-81.
21. Sargison ND, Jackson F, Bartley DJ, coll. Failure of moxidectin to control benzimidazole-, levamisole- and ivermectin-resistant *Teladorsagia circumcinda* in a sheep flock. Vet Rec 2005;156:105-9.
22. Sidani C, Astruc JM, Baelden M, coll. The French ovine scrapie plan: results and prospects. *In:* 9th World Congress for Genetics Applied to Livestock Production, Leipzig, 2010.

● **Élément indispensable de sélection durable, la diversité génétique apparait donc primordiale pour la résistance aux maladies.** C'est notamment ce que met en avant un avis de l'Anses, qui propose de maintenir une diversité génétique suffisante au gène PrP dans la population ovine française [1].

● **Une des possibilités pourraient être de ne sélectionner spécifiquement qu'un noyau limité de mâles pour la résistance aux maladies.** La diffusion de la semence des mâles de ce noyau serait alors, par exemple, destinée spécifiquement aux troupeaux se trouvant dans des zones de forte prévalence d'une maladie donnée.

● La sélection génomique considère les effets de tous les gènes de manière statistique sans connaissance des gènes impliqués dans le contrôle du caractère considéré. Étant donné les enjeux que l'on vient d'envisager, il serait souhaitable d'identifier précisément les gènes contrôlant la sensibilité à une maladie avant d'en faire un usage en sélection.

LA PLACE DU VÉTÉRINAIRE

● Quelle que soit l'option envisagée, sélection génomique ou sélection assistée par gènes, il convient d'être particulièrement prudent lors de la sélection pour la résistance aux maladies. La sélection doit être équilibrée après avoir étudié les relations entre les caractères de production et de résistance. Les priorités seront à adapter en fonction de la situation épidémiologique et les avis de vétérinaires seront indispensables.

● L'expertise des pathologistes est nécessaire à la définition de critères pertinents de mesure de la résistance aux maladies.
En effet, les bénéfices de la sélection génétique seront maximisés par la quantité et la qualité des informations traduisant la sensibilité des individus.

● **Dans cette optique, il sera indispensable d'élaborer, de standardiser et de généraliser l'enregistrement des données sanitaires individuelles de chaque animal.**

- **En Europe, l'exemple est donné par la Norvège** et l'élevage bovin laitier, où chaque animal dispose d'un "carnet de santé" dont les données sont transférées chaque mois au système d'informations national [17].
- **En France, l'identification électronique des petits ruminants,** désormais rendue obligatoire (règlement (CE) n° 21/2004) **constituerait un atout majeur pour la création d'une telle base de données.**

● De plus, le carnet sanitaire propre à chaque élevage pourrait permettre de tels enregistrements et ferait office d'interface entre vétérinaires, éleveurs et sélectionneurs. Il reste cependant à s'assurer de son bon usage et à analyser de façon pertinente les données qui y sont saisies. À ce titre, un rapprochement, dont les modalités restent entièrement à définir, entre les vétérinaires et les organismes de sélection sera des plus profitables.

CONCLUSION

● Il est probable que la génomique bouleversera l'élevage des petits ruminants, soit directement, par le passage à une sélection génomique ou assistée par gènes, soit indirectement, en permettant des avancées notables dans la compréhension du déterminisme génétique des caractères.

● Les opportunités offertes dans le cadre de la santé des petits ruminants sont nombreuses et s'inscrivent parfaitement dans le cadre d'un élevage plus respectueux de l'environnement et du consommateur. En offrant la possibilité de limiter le recours aux traitements pharmacologiques classiques, elle permet de prévenir l'apparition de pathogènes résistants aux molécules d'usage en médecine vétérinaire, mais aussi de limiter les résidus pharmacologiques dans les denrées d'origine animale. Cependant, la mise en place effective de ces schémas sera fortement dépendante du ratio bénéfices / coûts généré. Par ailleurs, la génétique s'inscrit dans un plan de maîtrise sanitaire à long terme, elle est donc complémentaire des stratégies vétérinaires classiques. ❐

formation continue

1. La valeur génétique d'un bélier Manech Tête Noire peut être estimée à partir des mêmes équations prédictives que celles des Lacaunes : ❑ oui ❑ non

2. La sélection d'animaux résistants sera plus efficace et plus durable en sélectionnant des centaines de gènes à effet faible plutôt qu'un seul gène à effet fort : ❑ oui ❑ non

3. L'atout majeur de la sélection génomique est la précision de ses estimations : ❑ oui ❑ non

RUMINANTS

NÉVA
EUROPARC 15, rue E. Le Corbusier
94035 CRÉTEIL CEDEX
Tél : (33) 1-41-94-51-51
Courriel : neva@neva.fr
NOUVELLES ÉDITIONS VÉTÉRINAIRES ET ALIMENTAIRES

Part III

Materials and methods

Cette deuxième partie consiste en une description de la souche du parasite utilisée dans les infestations expérimentales et des races ovines utilisées pour la création de la population hôte expérimentale. L'historique de la mise en place du protocole est rappelé, et les techniques de mesure de l'infestation sont décrites brièvement. Le traitement des données de génotypage est également rappelé.

In this section, descriptions of parasite strain and experimental animals are given. History of the experimental design has been recalled and measures of resistance are briefly described. Processing of genotyping data is also described.

5.8 *Haemonchus contortus*

H. contortus is one of the most pathogenic GIN of sheep. However it is not the most preponderant species in northern Europe nor in France where *T. circumcincta* and *T. colubriformis* are usually preponderant species [395, 80]. Still it has been used as a reference GIN for INRA works on host-pathogen interaction [288, 507]. More generally, *Haemonchus* has also been considered as a model for population genetics aspects in the study of GIN resistance to anthelmintics [185]. Its main biological features are recalled while justifications for its use are provided.

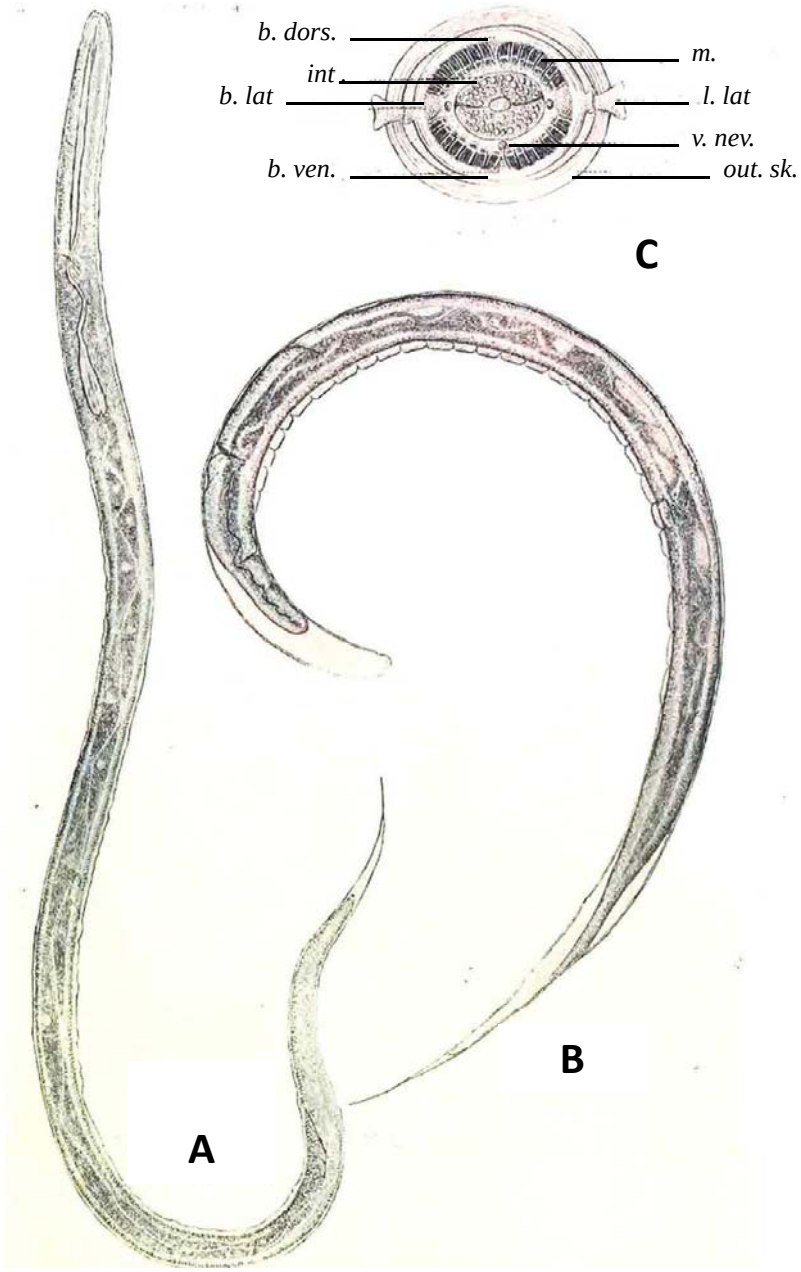
5.8.1 Morphology and life-cycle

Pioneer description of the anatomy and life-cycle of *H. contortus* has been performed by Veglia in 1915 [529]. This experimental work aimed at an exhaustive description of the morphology and the impact of environmental factors on each parasitic stage, from eggs to adult males and females [529]. A brief overview of *H. contortus* characteristics will be reminded in this paragraph based on this work [529].

As a trichostrongylid, *H. contortus* evolves following a direct cycle split into a free life period on the pasture and a parasitic life within the host (see section 1.2 and figure 1.1). Eggs produced by the females are released on pasture where they hatch into first-stage larvae (L1) within approximately one day according to environmental conditions before subsequently moulting into second-stage larvae (L2). Finally L2 develop into the infective third larval stage that wait for uptake by its host, sealed off in its cuticle (see figure 5.4). According to optimal moisture and temperature conditions, the developmental process can be achieved within seven to ten days. Each larval stage can be divided into two distinct periods, the first one being dedicated to feeding and growing in preparation of the second step in which the larvae are rather lethargic and undergo structural changes.

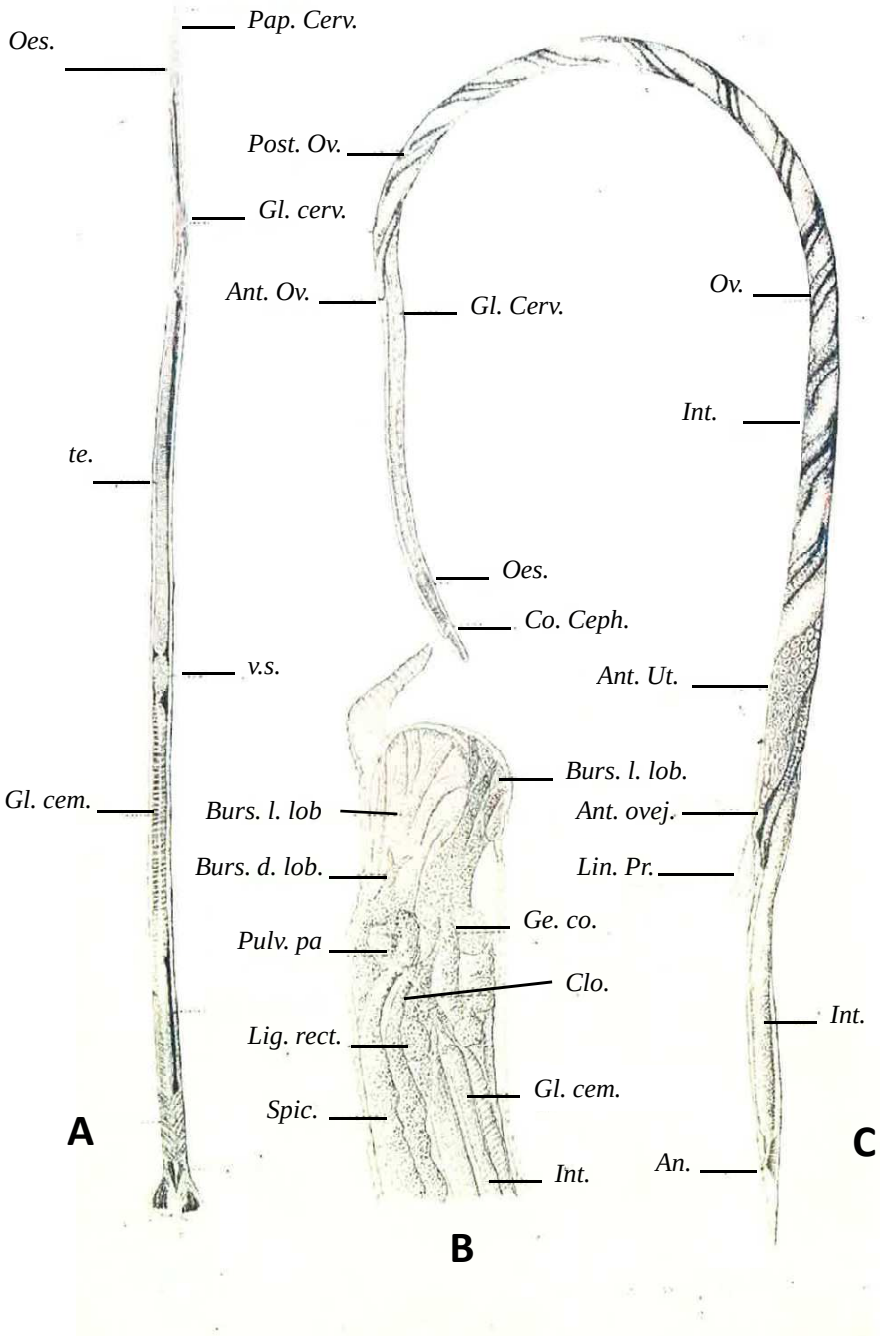
After being ingested by their host, L3 larvae end their free life and begin their parasitic life by getting out their cuticle once they reach the rumen and begin to feed without piercing the abomasal mucosa. Within one to four days after entering the host, L3 have evolved into L4 stage that attaches itself to the mucosa hence resulting in mucosal damages. This stage lasts for about one week before moulting into the immature adult (L5) and adult worm, males being somewhat smaller than females (size of 5 mm and 6.5 to 7.9 mm just after ecdysis for males and females respectively, see figure 5.5).

Figure 5.4: Morphology of the mature third stage larvae



A: Mature larva, after about a week of maturity. B: Mature larva two months old and still alive. C: Transversal section of a mature larva, made in the anterior portion of the chyle intestine. *b.dors*: dorsal band, *b.lat.*: lateral band, *b.vent.*: ventral band, *int.*: intestine, *m.*: muscle, *v. nev.*: ventral nev., *out. sk.*: outter skin. Reproduced from Veglia (1915) [529]

Figure 5.5: Morphology of the *H. contortus* male and female



A: Adult male. B: Posterior end of a male in 4^{th} ecdysis. C: Adult female. *oes.*: oesophagus, *pap.cerv.*: cervical papilla, *Gl. cerv.*: cervical glands, *int.*: intestine, *ov.*: ovary, *te.*: tegument, *co.ceph.*: cephalic commissure, *v.s.*: seminal vesicle, *burs. l. lob.*: lateral lobe of the bursa, *burs. d. lob.*: dorsal lobe of the bursa, *ge. co.*: genital cone, *gl. cem.*: cement gland, *pulv. pa*: pulvillum postnatalis, *lig. rect.*: rectal ligament, *clo*: cloaca.Reproduced from Veglia (1915) [529]

Hence, under optimal condition, the pre-patent period lasts approximately 21 days. Otherwise a phenomenon of arrested development or "hypobiosis" can take place, *i.e.* the withholding of a large numbers of larvae at the same stage during a time period longer than required to reach that stage [505]. Even if the nature of the stimulus triggering hypobiosis is still controversial, it provides a good way to survive adverse climatic conditions [515]. Interestingly, Jacquiet *et al.* demonstrated that an increased lifespan of adult worms for more than 50 weeks over the dry season rather than an hypobiotic period, resulted in the completion of life cycle from one wet season to another [249]. Both phenomena underline both the evolutionary capacity of *H. contortus* to its parasitic life and the variations in this parasitic life style.

5.8.2 Justification

For experimental purposes, a challenge with controlled infection dosis should be performed. In addition, isolating genetic components of the immune response to GIN is also synonymous of being able to split individuals in distinct groups corresponding to resistant or more susceptible individuals. In addition, the GIN species used for such experimental infection should be easy to maintain.

Thus *H. contortus* has been chosen as the reference nematode for studying GIN infection in sheep. As already described, *Haemonchus* females are highly fecund and larvae are easily cultured hence favoring the production of infection doses and making FEC a good proxy for worm burden. In addition, *Haemonchus* is highly pathogenic in naive animals hence providing distinct pattern between more resistant and more susceptible lambs, especially with an infection dosis of 10,000 larvae.

As genetic correlation between resistance to this genus and to other trichostrongylids has been estimated to be almost equal to one [199], transposing findings about genetic resistance to *H. contortus* to other trichostrongylid species might be relatively straightforward.

5.9 Exploiting the Romane and Martinik Black Belly breeds in a back-cross population

5.9.1 Breeds' history and description

5.9.1.1 The Martinik Black Belly breed

The French West Indies harbor very similar sheep populations originating from Africa, like Barbados Black Belly, West African or Virgin Island White. These different group show similar traits as other breeds encountered in the Caribbean, like St Croix, Gulf Coast Native, *i.e.* hairy hardy breed well adapted to tropical breeding conditions. In 1993, all the different genetic groups from the French West Indies were gathered into a single ***Martinik*** breed in 1993. Martinik animals are hairy and polled with dark-colored mucosa. Due to this mixture of different ovine populations, the Martinik breed can be categorized into four types showing different coat patterns :

- Créole, with dark grey coat
- Black Belly, with red coat and black belly and limbs
- Saint Martin, with red hair and white or red belly and limbs
- Blanc, totally white

They exhibit long and fine limbs with large breast and back. The selection nucleus relies on nine farms harboring 1,000 controlled ewes. Breeding objectives focus on maintaining both the breed adaptation to the grazing tropical system and the breed reproduction performances. In addition there is a will to increase maternal performances like milk production and lambs growth.

The Martinik Black Belly, denoted MBB, is one of the pure breed contributing to the back-cross population used in this work.

5.9.1.2 The Romane breed

The Romane breed (denoted RMN) is a synthetic breed created in 1980 by INRA geneticists. The idea was to benefit both breeds qualities, *i.e.* meat production conferred by the Berrichon du cher breed and the high prolificacy of the Romanov breed. After four generations of crossing,

the RMN breed was born. The RMN breed is particularly easy to manage and shows good maternal capacities. RMN animals are usually white but black sheep or black spotted animals regularly pop up to recall the Romanov origin. The RMN population counts more than 45,000 individuals with 500 AI rams.

5.9.2 Breeds' performances

Main features about breed performances have been summarized in table 5.5 (data obtained from the Bureau des Ressources Génétique database, www.brg.prd.fr). Briefly, the RMN breed is one of the most prolific breeds in France and lamb growth is about two fold in comparison to the MBB breed. However this latter breed also performs well in terms of reproduction with almost two lambs born at lambing and a reduced interval between successive lambings.

Table 5.5: Mean production statistics of the Romane (RMN) and the Martinik Black Belly (MBB) breeds

Trait	**Breed**	
	MBB	**RMN**
Age at first lambing *(in days)*	420	420
Prolificacy *(No. lambs/lambing)*	1.8	2.6
Between lambing interval *(in days)*	147	300
Weight daily gain before weaning *(in g/day)* [a]	180 - 167	285-266
Age at weaning *(in days)*	90	60
Weight at weaning *(in kg)* [a]	11-9	25-21
Adult weight *(in kg)* [a]	65-40	95-75

a: sire value-dam value

5.9.3 Immuno-pathological comparison of the two breeds under infection

Early works by Yazwinski *et al.* have shown the greater resistance potential of the Black Belly breed towards *H. contortus* infection in comparison to the Dorset breed [571].

In the early 2000's, Aumont *et al.* and Gruner *et al.* compared the different reaction of MBB and RMN during *H. contortus* challenge. Aumont *et al.* first started by comparing each breed in its own environment, *i.e.* MBB animals in Guadeloupe and RMN in France. Each breed was infected with a sympatric and an allopatric strain isolated from goats in Guadeloupe or in France and primed and naive lambs were also compared [24]. The strain comparison showed that the Guadeloupe strain was significantly more successful in establishing in its host as 1.5 fold FEC

increase was reported in comparison to the other strain. No matter what strain was used, the MBB animals performed better than their RMN counterparts. More generally, the MBB breed was demonstrated to better cope with *H.contortus* in comparison to the RMN breed. Indeed FEC were always significantly lower in the MBB breed across all comparisons performed, while worm burden never reached 2,000 worms [24]. The worms establishment rate was three fold more important in the naive RMN lambs and this difference stretched to an 8 fold difference between primed lambs of each breed even if RMN primed lambs performed better than naive lambs at reinfection. Interestingly the eosinophil counts remained constant in the naive RMN lambs whereas an increase was reported for all other groups, *i.e.* naive and primed MBB and primed RMN.

After this first experiment, Gruner *et al.* set up another trial to confirm these differences under temperate conditions in the French Région Centre and to investigate the performances of MBB (n=40), RMN (n=88) and their F1 cross-products (n=84) towards *T. colubriformis* and *T. circumcincta* infections. Not only confirming the better resistance of the MBB individuals over the RMN lambs under temperate conditions, the authors demonstrated that this resistance was also observed for an infection by *T. colubriformis*. No results could be obtained for *T. circumcincta* as not enough MBB individuals were available. However the authors also found that the F1 lambs outperformed the RMN lambs after a primary challenge whatever the tested nematode species hence fitting the MBB pattern [198].

One year later, the heritability of the resistance to *H.contortus* was published in a RMN flock of 150 lambs from 30 families [199]. Lambs were infected twice with a one week drenching period and with 10,000 larvae of *H.contortus* before being infected two months later with a 10,000 *T.colubriformis* dose and conversely. Genetic parameters were in the usual range of values, *i.e.* 0.39 and 0.48 for *H.contortus* and somewhat lower for *T.colubriformis*. In addition the authors reported genetic correlations very close to one between the two species suggesting a non-specific resistance mechanism being independent of the first encountered nematode [199]. In addition they also found this type of correlation between the first and second infection suggesting that resistance is rapidly acquired. Even if well characterized for the RMN breed, no genetic parameters have been published for the resistance to GIN in the MBB breed. However results of experimental infection also show some inter-individual variation that might translate the underlying genetic variation.

Subsequently, Terefe *et al.* aimed at deciphering the underlying immune mechanisms of

the response to *H.contortus* in the two breeds, and especially focused on the role played by eosinophils during the infection course. In a first paper, they confirmed the previous results by Lacroux *et al.* [288] that *H.contortus* elicited a Th2-biased immune response in both breed. They also demonstrated that this Th-2 response was more pronounced in the resistant breed at first infection, as the IL-4, IL-5 and IL-13 expressions were up-regulated in the MBB [507]. Such an efficient response appeared at reinfection in the RMN breed. During this infection, the authors also highlighted a higher blood eosinophilia in the resistant breed and demonstrated in an *in vitro* study that these cells could reduced the establishment potential of infective larvae [506]. Subsequently, the putative functional differences between isolated eosinophils from the resistant and susceptible breeds were investigated but no significant differences in the larval immobilizing ability could be observed even if the MBB breed exhibited very high tissue eosinophilia in comparison to the RMN breed [508].

The genetic variation within the RMN breed has been quantified and some functional explanations have been proposed to explain between breeds differences. However, the genetic basis of these differences have been partially resolved. Especially, no link between these different approaches,*i.e.* where and what are the genes responsible for those differences, have been investigated.

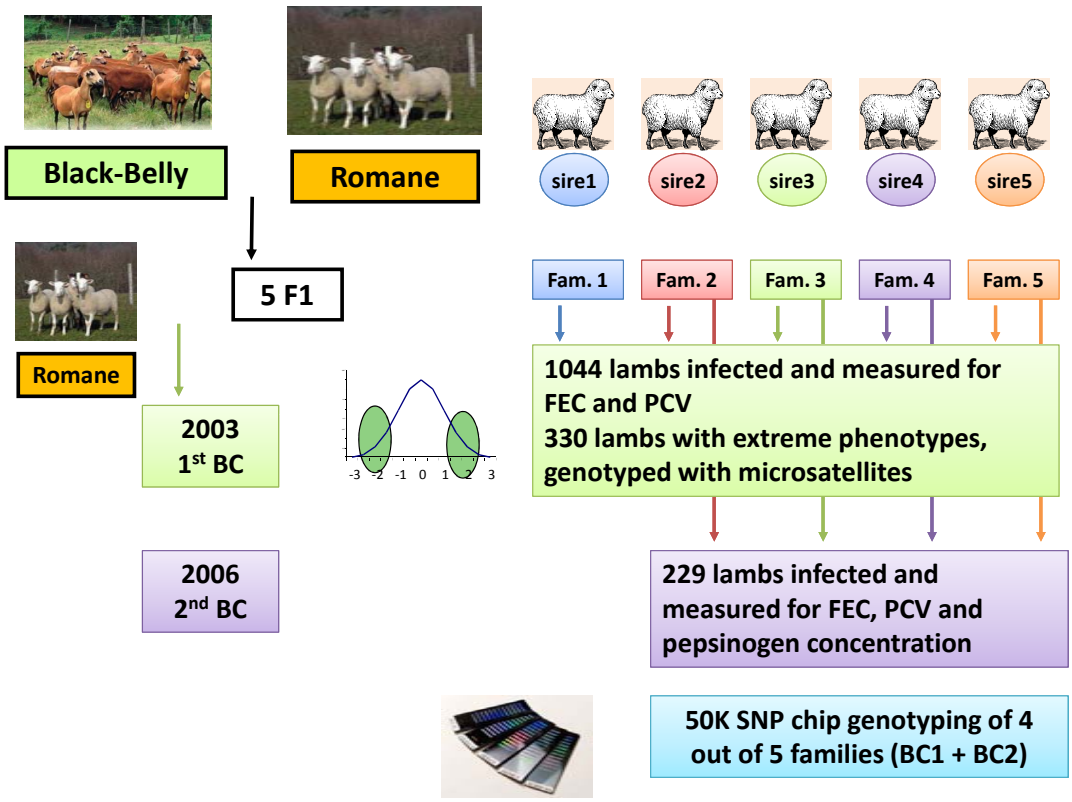
5.9.4 Back-cross (BC) history (see figure 5.6)

To screen for QTL affecting resistance to nematodes, a BC population was created at the INRA La Sapinière experimental farm. Five F1 sires were produced by mating five MBB sires to five RMN ewes. F1 sires were subsequently back-crossed to RMN ewes that gave birth to 1044 BC lambs (BC1). Among these lambs, 330 individuals with extreme FEC were genotyped in a so called "selective genotyping" approach to minimize genotyping costs while increasing the probability of finding a QTL (see section 5.11).

After two successive challenges with *H. contortus*, lambs were slaughtered for fine phenotyping, *e.g.* parasitological measurements, anti-Haemonchus IgG.

In 2006, four out of the five original F1 sires were mated again to other RMN ewes to produce 229 BC2 lambs that were trickled twice and measured for FEC, PCV and pepsinogen concentration before infection and at 15 dpc.

Figure 5.6: Schematic representation of the BC design implemented



BC: back-cross

5.10 Experimental infection and phenotyping

5.10.1 Infectious strain and infection procedure

The *H.contortus* strain used in this work, called the "Humeau" strain has been isolated from goats in southwestern France. It has been regularly multiplied on experimental Tarasconnais lambs. Faeces from these lambs are collected before being stored at 4C. Before infection, faeces containing larvae are put into water and exposed to day light to collect alive larvae only (Baerman procedure). For experimentation purposes, naive lambs under study are dosed with 10,000 infective larvae.

5.10.2 Fecal Egg Count, FEC

FEC are measured according to the McMaster technic modified by Raynaud [428]. Briefly, three grams of faeces taken from the rectum are crushed and diluted in 42 mL of saturated saline water (200g NaCl in 600 ml H2O: density =1.8). After one filtration, two aliquots are sampled and put into the two chambers of a McMaster slide, each containing 0.15 mL. In the end, 0.3 mL from the 45 mL total volume and the obtained count is then multiplied by 50 to obtain the total number of eggs per gram of faeces.

5.10.3 Haematological parameters

L4 stages and adult worms feed blood from their host. Therefore packed-cell volume correlates to the worm burden. However the resulting PCV value is a mixture of both the parasite uptake and the host regeneration hence providing an indirect measure of the host resilience.

Packed-cell volume was determined as follows. A blood sample, taken in EDTA coated tubes, is put into capillary tubes that are subsequently centrifuged at 12000 rpm for 10 minutes. After centrifugation, hematocrit value is determined by measuring the length of the red blood cell on a reference grid.

Other hematological examination were performed using the Sysmex XT-2000iV hematology analyzer at the Ecole Nationale Vétérinaire de Toulouse (C. Trumel and A. Geffré). This automatic analyzer gives access to counts of red and white blood cells populations as well as platelet number and hemoglobin concentration parameters.

5.10.4 Parasitological data: worm burden, worm female fertility

At the end of infection, some lambs were sacrificied by intravenous injection of 6 mg/kg embutramide. The abomasum was opened and its contents and washings were collected and passed through a 40 μm sieve. Remaining content was put into absolute alcohol for storage before subsequent worm counting. The abomasum was then digested in pepsin-hydrochloric acid solution (37C, 6h) to collect the tissue-dwelling worms [224]. The solution was prepared by mixing 20 gm of pepsin, 20 ml of HCl and distilled water to make a final volume of 1 litre. After digestion, solutions containing the remaining larval and adult stages were collected and preserved in absolute alcohol.

The volume of contents and abomasal digestion was adjusted to 2 litres and 10% aliquot were sampled for worm burden determination, worms being classified as adult male and female, immature male and female or L4 stages.

In addition, to the total worm burden, female worm fecundity was determined by measuring the length of 35 females per lamb. A second measure was based on the number of eggs *in utero* determined for 20 randomly chosen female worms per animal. Each intact female was put into a mixture of 40 mL of mild bleaching agent (Milton Sterilizing fluid containing 2% w/v sodium hypochlorite and 16% w/v sodium chloride Milton Pharmaceutical LTD) diluted in 160 mL of distilled water [505], and all eggs liberated from the uterus were counted.

5.10.5 Pepsinogen concentration

Under physiological conditions, pepsinogen is transformed into pepsin, a gastric enzyme. Under abomasal infection, *e.g. O. ostertagi* in cattle or *H. contortus* in sheep, mucosal damages prevent the cleavage of pepsinogen into pepsin hence resulting in an increased pepsinogen concentration in plasma [164, 471]. Hence pepsinogen can be used as a biomarker of infection [87].

Pepsinogen concentration was hence determined using the micro-routine determination proposed by Dorny and Vercruysse [129] before infection and 15 days after infection for both QTL mapping and functional validation.

Briefly, the serum sample was acidified with HCl and incubated overnight at 37C with bovine serum albumin (BSA) before being stopped with 4% trichloro-acetic acid (TCA). The resulting mixture was centrifuged at 14000 rpm for 5 minutes. An aliquot from the supernatant was added to 0.25 M NaOH and the plates were incubated at room temperature with folin reagent for 30 minutes. The liberated tyrosine was estimated by reading the absorbance at 680 nm and

the values were expressed as unit (U) or milliunit (mU) tyrosine/litre of serum.

5.10.6 RNA extraction and cDNA amplification

Total RNA from abomasal fundic mucosa and draining lymph nodes of the sampled animals were extracted following the commercial RNeasy Mini Kit (Qiagen). The RNA quality of the recovered RNA was monitored by A260/A280 spectrophotometry. RNA were subsequently reverse transcripted to cDNA with a Reverse Transcriptase commercial kit (Invitrogen).

5.10.7 Gene expression study

In a previous study (Genohpar project, [316]), some genes had been found differentially expressed between the MBB and the RMN breed, either in abomasal mucosa (GAL15, ITLN2, TFF3) or in draining lymph nodes (OX40, CXCL14, CCL16) or both (IL4, IL13, $TNF\alpha$, $IFN\beta$). The relative expression of each of these genes was therefore tested in BCxBC lambs.

In addition, the annotated genes the closest to QTL regions of interest were retrieved from the second assembly of the ovine genome (http://www.livestockgenomics.csiro.au/cgi-bin/gbrowse/oarv2.0/).

Primers were designed for these particular genes using the "primer 3" NCBI website and the bovine transcript as a template (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) . Secondary structures were looked for on the Mfold website (mfold.rna.albany.edu, [581]) and selected primer sequences were blasted against the ovine genome to ensure specificity of their target.

The qPCR was performed with three replicates per sample. A set of four reference genes specific of each tissue were determined according to their gene-wise stability value as reported in [527]. Differential expression was tested following the DDCt method [319]. The cycle time (Ct) value of the gene of interest was corrected by the average level of reference genes expression. Ct values of the infected animals were corrected by the average Ct value of the corresponding control animals. Subsequently a Wilcoxon test was applied to determine any significant difference between the compared groups, *i.e.* $p < 0.05$. The complete data processing was performed using an homemade R script (R software, http://CRAN.R-project.org/doc/FAQ/R-FAQ.html).

5.11 Genotyping strategy

5.11.1 Selective genotyping approach with microsatellites

The QTL mapping experiment originally started in early 2000's when no dense SNP chip were available but microsatellites. These markers are labour intensive and expensive to develop and to genotype, therefore a limited panel was used.

To avoid hampering too much the QTL detection analysis, a so called "selective genotyping" procedure was attempted.

Hence 330 animals with extreme FEC were chosen.

5.11.2 Entering the genomic era: processing of the SheepSNPQTL project SNP data

Soon after the release of the 50K SNP chip, a French research project started that aimed at mining the genetic basis of complex traits of interest [375]. Main features of the 50K SNP chip and the SSQ project have been summarized in this section and the quality checking applied to the SNP data are described. Thanks to the sequencing effort of the International Sheep Genomics Consortium, 50,000 high-quality SNPs have been detected in 74 sheep populations and pooled together on a DNA SNP chip. In 2009, this SNP chip was released by Illumina.

5.11.2.1 The SheepSNPQTL project

The SheepSNPQTL project aimed at taking advantage of the information provided by the 50K available SNPs in French sheep populations to both mine traits of interests and try to apply genomic selection to the Lacaune dairy breed [375]. Around 4,000 animals were genotyped to study five traits of interest, covering resistance to disease, production and animal behaviour (see table 5.6).

Table 5.6: Available populations in the SheepSNPQTL project [375]

Trait under study	Breed	Experimental design	Population size
Resistance to GIN	RMN*MBB	Back-cross	1192
Milk production	Lacaune, MTR	Grand-daughter	1293
Behaviour, meat production	RMN	Sire	1095
Resistance to mastitis	Lacaune	Case-control	305
Resistance to scrapie	Lacaune	Case-control	61

Key: *RMN*: romane, *MBB*: Martinik Black Belly, *MTR*: Manech Tête Rousse

5.11.2.2 SNP quality check [453]

As a part of this PhD project, SNP data of the SheepSNPQTL project were processed altogether for quality check before being uploaded on the SIGENAE storage platform.

First SNP editings were applied to get rid of useless SNP in every population, *e.g.* ungenotyped SNP and markers monomorphic in every population.

Thanks to the CSIRO assembly effort, a second version of the ovine genome was released in late 2010. In this version, physical positions of 53,648 SNPs out of the 54,955 original SNPs have been updated. Two additional groups of SNPs were also present, *i.e.* 427 with two plausible positions and 593 without position. According to B. Dalrymple in charge of the ovine genome assembly, a significant proportion of the SNPs that were not included in the new assembly were in regions with high GC content and scattered across the genome rather than clustered. He sugested the positions of SNPs that were not included in the second version of the assembly were as likely as the ones of included SNPs, so that unknown positions should be interpolated for not included SNPs rather than just discarding them. The interpolation was based on the SNPs flanking the marker not included. Let's consider M2, the SNP not included in the second assembly and M1 and M3 its flanking marker. If M1 and M3 had been repositioned in the second assembly, then M2 was placed at the same relative position as in the first assembly following :

$$M2' = M1' + (M3'–M1') * (M2 - M1)/(M3 - M1)$$

with $M1', M2', M3'$ being the positions of the second assembly.

In the end, 404 SNPs had their position inferred.

A data workflow was implemented to process the file produced by the LABOGENA genotyping platform, perfom SNP editing and to deliver clean datasets (see figure 5.7. SNP are genotyped using a fluorescence based method that results in a file clusterizing homozygotes and heterozygotes animals (figure 5.8). The more distinct are the three genotypic groups, the better is the SNP call rate. A file summarizing every luminescence signal was produced by LABOGENA and was subsequently processed in our lab for creating a .csv file (A. Tircazes).

A fortran coded program was kindly provided by A. Ricard for checking SNP data. This original script was upgraded to perform verifications in several sheep populations. This program performs a first read of the genotype file to count the number of SNP genotyped per individual and the allelic frequencies of every SNP in each considered population. A second round of

Figure 5.7: SNP data workflow

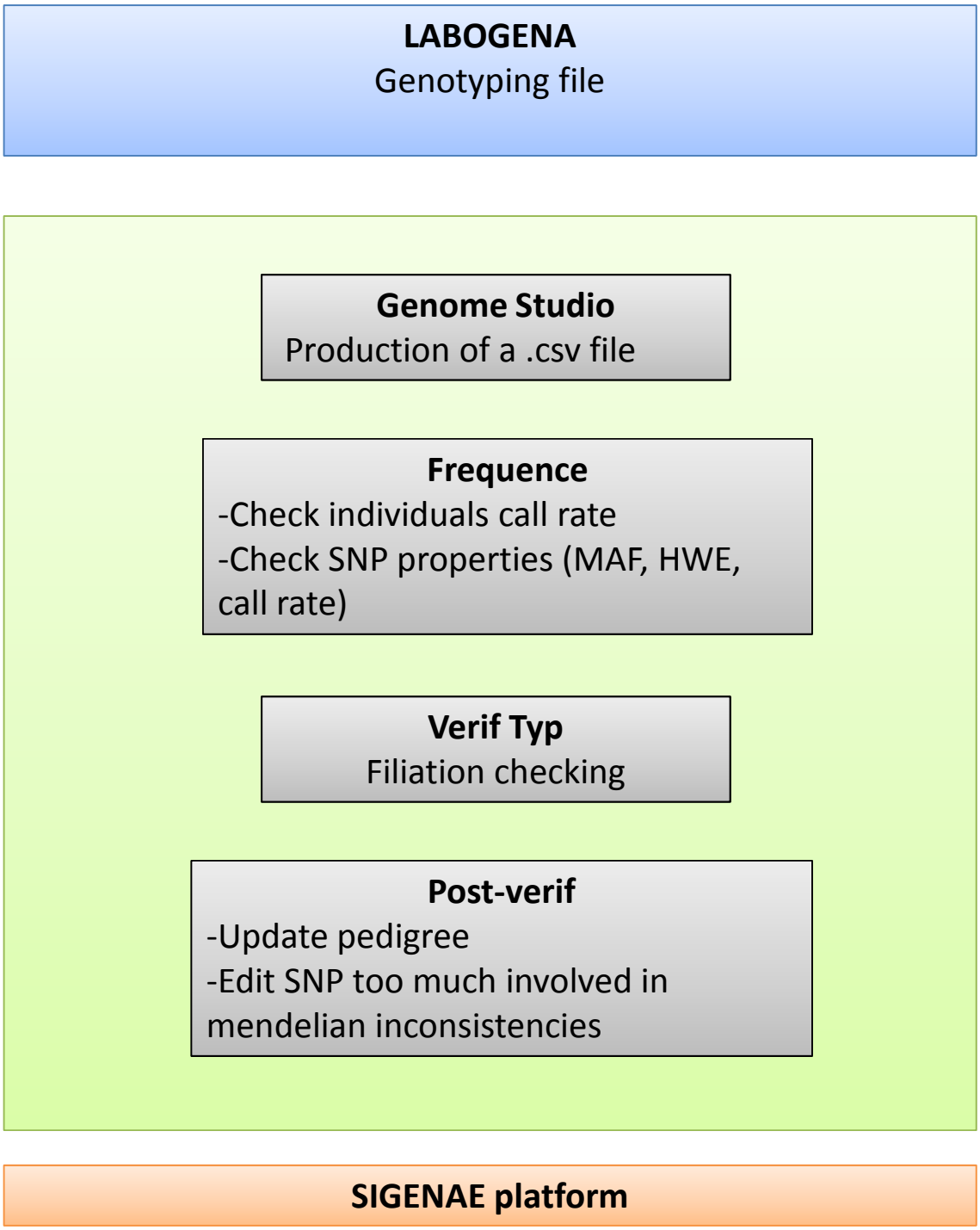
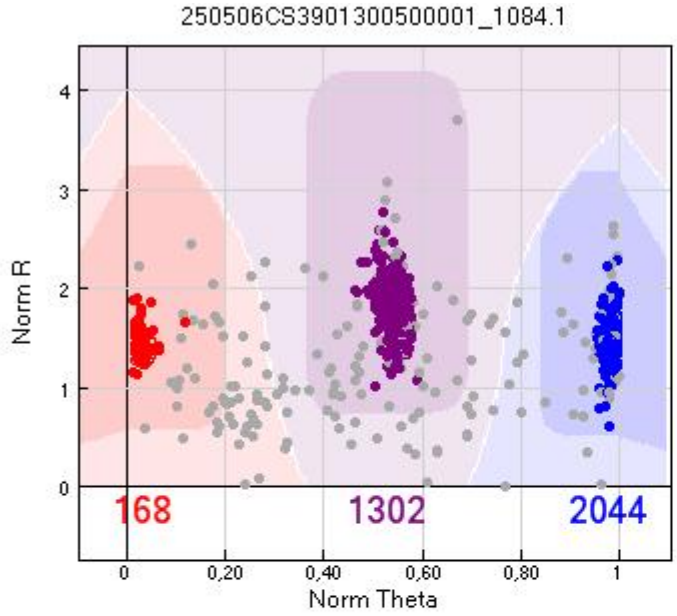


Figure 5.8: Fluorescence cluster



reading is subsequently performed to discard animals with poor genotype prediction (call rate below 98% of the useful SNPs).

Soon after the release of the ovine SNP chip, the HAPMAP consortium released a list of SNPs with abnormal behavior that should not be considered for subsequent analyses (J. Kijas, personal communication):

- SNPs that Illumina annotated as abnormal
- SNPs with minor allele frequencies equal to zero
- SNPs that displayed discordant genotypes between experiments
- SNPs showing Mendelian inconsistencies within the International Mapping Flock

These SNPs were hence edited from our dataset as well.

SNPs failing a range of tests are also eliminated from the dataset. Considered quality parameters were those generally reported in literature [555]. Hence, useless SNPs (monomorphic or ungenotyped in the whole dataset) and any SNP with a call rate less than 97% were discarded. In addition, a test for Hardy-Weinberg disequilibrium was performed and a conservative threshold of $p < 10^{-6}$ was applied to get rid of SNP with abnormal behaviour while letting each data user deciding for more restrictive threshold to be applied. In addition, a Minor Allele Frequency of 1% was applied. Due to the particular structure of some experimental populations (case/control, back-cross populations) these two latter criteria were only considered in the Lacaune and Romane populations.

After first editing of the SNP data, detectable genotyping errors were corrected by checking any mendelian inconsistencies between parents and offspring using the VerifTyp program provided by F. Guillaume (INRA, GABI). The purpose of this checking was two-fold, *i.e.* correcting detectable genotyping errors and detecting pedigree errors. After a first round of verification, the distribution of the total number of mendelian inconsistencies per individual was plotted and a threshold of 60 inconsistencies per individual was considered to determine pedigree errors. In the end 29 animals with incorrect genealogy were removed. In addition, any SNP occurring with too high a frequency in genotyping errors was discarded (n=34).

In the end, each project leader was subsequently in charge of completing the editing in his own population according to his preferred criteria and to the analysis run.

Part IV

Experimental work

Chapter 6

Exploitation of the ovine DNA SNP chip for mapping the QTL affecting resistance to *H.contortus*

Résumé

Les races MBB et RMN montrent des capacités de résistance à *H. contortus* très différentes, les agneaux MBB étant habituellement moins sensibles alors que les agneaux RMN peuvent atteindre des niveaux d'hématocrite critiques pour leur survie [507]. Tirant parti de cette variabilité, un troupeau BC a été créé à la ferme expérimentale de La Sapinière pour rechercher des QTL affectant la résistance à *H. contortus*.

La population complète a subi deux infestations successives par *H. contortus* et l'intensité d'oeufs excrétés et le micro-hématocrite ont été mesurés. Un sous-ensemble de 332 agneaux aux performances extrêmes a été autopsié pour du phénotypage fin.

En se basant sur les résultats d'une étude préliminaire utilisant des marqueurs microsatellites, quatre des cinq familles de père ont été génotypées avec la puce ovine 50K. Les données SNP ont contribué à augmenter la précision de la localisation des QTL et ont également permis de rechercher des QTL sur des chromosomes jusqu'alors délaissés par les études utilisant des microsatellites. Parmi les QTL identifiés, cinq régions sont sortis du lot sur les chromosomes 5, 7, 12, 13 et 21. Un QTL situé sur le chromosome 12 était particulièrement intéressant puisqu'une région limitée de 10 Mbp affectait l'intensité d'oeufs excrétés en première et deuxième infestation. Une portion limitée de OAR21 a également été associée à la variation de concentration en pepsinogène

et colocalise avec le gène PGA5 qui code pour le pepsinogène. Ces résultats sont présentés en intégralité dans le deuxième article (section 6.1.1).

Suite à cette analyse QTL, différentes méthodes QTL ont été comparées en utilisant l'information génomique du chromosome 12 génotypé dans la population back-cross et les mesures correspondantes d'intensité d'oeufs excrétés en première infestation. Les résultats préliminaires ont confirmé les résultats déjà obtenus avec QTLMAP, et des simulations devraient permettre de finaliser la comparaison des différentes méthodes.

De plus, des données génomiques additionnelles ont été collectées dans les races pures pour rechercher d'éventuelles traces de sélection apparues dans l'une des deux races, et qui pourrait éventuellement colocaliser avec un des QTL déjà identifiés. De manière intéressante, une trace de sélection de 5 SNPs a été identifiés dans la race MBB, les mêmes marqueurs ségrégeant toujours dans la population RMN. Ce sweep a été localisé à 42 Mbp, c'est-à-dire tout à coté de la position du maximum de vraisemblance du QTL associé à l'intensité d'oeufs excrétés chez les agneaux immuns.

Summary

The MBB and RMN breeds exhibit opposed phenotypes to infection by *H. contortus*, the MBB lambs usually showing reduced signs of infection while some RMN lambs can reach life-threatening levels of hematocrit [507]. Exploiting this genetic variability, a BC flock was created at the La Sapinière experimental unit to look for QTL affecting resistance to *H. contortus* infection that may explain the observed differences between the two breeds.

The whole population was challenged with *H. contortus* in two consecutive experimental infections and FEC and packed cell volumes were measured. A subgroup of 332 lambs with extreme FEC was sacrificed to measure additional parasitological and pathophysiological traits.

A preliminary study had been performed using 160 microsatellites markers before the release of the ovine DNA SNP chip in 2009. Four out of the five original families that segregated for most of the identified QTL were genotyped for the 50K SNP chip.

In the end, SNP data contributed to increase accuracy of mapping and help detecting QTL on chromosomes that had been poorly covered with microsatellites. Among the detected QTL, five outstanding regions were found on OAR5, 7, 12, 13 and 21. A QTL on OAR12 was of

particular interest as a 10 Mbp-wide region affected FEC at 1^{st} and 2^{nd} infection. On OAR21, a thin QTL region was associated to pepsinogen concentration and surrounded the *PGA5* locus that has been known to code for the pepsinogen protein. These results are fully presented in the second paper (section 6.1.1).

In addition to this QTL mapping experiment, performances of various QTL mapping methods have been compared using the OAR12 genomic data of the back-cross population and corresponding FEC at first infection. Preliminary results confirmed the published findings obtained with the QTLMAP software and simulations still need to be performed for a proper comparison of mapping methods.

Further, additional data were collected from pure breed populations to look for any selection sweep occurring in one of the two breeds that could eventually colocalize with one of the already identified QTL. Interestingly, a 5-SNP region was found to have reached fixation within the MBB breed but still segregated within the RMN population. This sweep region was mapped at 42 Mbp, *i.e.* close to the maximal LRT position of the QTL affecting FEC in immune lambs.

6.1 Second paper: QTL mapping study

6.1.1 Paper No. 2

American Society of Animal Science

Publisher of the *Journal of Animal Science*
http://www.asas.org

June 29, 2012

To whom it may concern:

The article entitled, “A genome scan for QTL affecting resistance to Haemonchus contortus in sheep”, with Guillaume Sallé as corresponding author was accepted for publication in the *Journal of Animal Science* and will be published in the August or September 2012 issue.

The paper will be in volume 90 and the page number will be assigned approximately 30 days before publication.

If there any questions, do not hesitate to contact me at steven.zinn@uconn.edu.

Sincerely yours,

Steven Zinn
Editor-in-Chief
Journal of Animal Science

A genome scan for QTL affecting resistance to *Haemonchus contortus* in sheep[1]

**G. Sallé,*†[2] P. Jacquiet,† L. Gruner,‡ J. Cortet,‡ C. Sauvé,‡, F. Prévot,†
C. Grisez,† J. P. Bergeaud,† L. Schibler,§ A. Tircazes,* D. François,* C. Pery,# F. Bouvier,#
J. C. Thouly,# J. C. Brunel,# A. Legarra,* J. M. Elsen,* J. Bouix,* R. Rupp,* and C. R. Moreno***

*INRA, UR631, Station d'Amélioration Génétique des Animaux, BP 27, F-31326, Castanet-Tolosan, France; †INRA, UMR1225, Interactions Hôtes—Agents Pathogènes, BP 87614, F-31076 Toulouse, France; ‡INRA, UR1282, Infectiologie Animale et Santé Publique, F-37880 Nouzilly, France; §INRA, UMR1313, Laboratoire de Génétique Biochimique et de Cytogénétique, F-78252 Jouy-en-Josas, France; and #INRA, UE332, Domaine de la Sapinière, F-18390, Osmoy, France.

ABSTRACT: Gastrointestinal nematodes are one of the main health issues in sheep breeding. To identify loci affecting the resistance to *Haemonchus contortus*, a genome scan was carried out using 1,275 Romane × Martinik Black Belly backcross lambs. The entire population was challenged with *Haemonchus contortus* in 2 consecutive experimental infections, and fecal egg counts (FEC) and packed cell volumes were measured. A subgroup of 332 lambs with extreme FEC was necropsied to determine the total worm burden, length of female worms, sex ratio in the worm population, abomasal pH, and serum and mucosal G immunoglobulins (IgG) responses. Pepsinogen concentration was measured in another subset of 229 lambs. For QTL detection, 160 microsatellite markers were used as well as the Illumina OvineSNP50 BeadChip that provided 42,469 SNP markers after quality control. Linkage, association, and joint linkage and association analyses were performed with the QTLMAP software. Linkage disequilibrium (LD) was estimated within each pure breed, and association analyses were carried out either considering or not the breed origin of the haplotypes. Four QTL regions on **[AU: define OAR]** OAR5, 12, 13, and 21 were identified as key players among many other QTL with small to moderate effects. A QTL on OAR21 affecting pepsinogen concentration exactly matched the pepsinogen (*PGA5*) locus. A 10-Mbp region affecting FEC after the 1st and 2nd infections was found on OAR12. The SNP markers outperformed microsatellites in the linkage analysis. Taking advantage of the LD helped to refine the locations of the QTL mapped on OAR5 and 13.

Key words: *Haemonchus contortus*, linkage disequilibrium, quantitative trait loci, resistance, sheep

J. Anim. Sci. 2012.90:1–16
doi:10.2527/jas2012-5121

INTRODUCTION

Gastrointestinal nematodes (**GIN**) are one of the main health issues in grazing ruminants (Coop et al., 1985; Mandonnet et al., 2003; Mandonnet et al., 2005; Davies et al., 2006). Among GIN, *Haemonchus contortus* settles in the sheep abomasum where it sucks blood from its host, thus leading to anemia and even to death in the absence of treatment. This issue has worsened over recent years because GIN have developed resistance against anthelmintics, the usual means of control. Resistance is constantly increasing in terms of prevalence, geographical repartition, and severity (Sargison et al., 2007; Traversa et al., 2007; Howell et al., 2008; Hoglund et al., 2009; Cezar et al., 2010).

Selection for animals that could resist nematode infection may provide a feasible long-term control strategy (Bishop and Morris, 2007). Furthermore, subtropical and tropical sheep breeds, such as Gulf Coast Native, St Croix, Santa Ines, or Barbados Blackbelly are considered to be relatively resistant to *H. contortus* (Aumont et al., 2003; Amarante et al., 2004), hence contrasting with other breeds from temperate areas, like

[1]The authors would like to thank the staff of the INRA experimental farm La Sapinière (Osmoy, France) for managing the experimental flock as well as for sampling all lambs. The technical staff of the Nouzilly experimental unit is acknowledged for handling lambs at the INRA abattoir as well as S. Normand for his participation in the analysis of parasitological samples. The QTLMAP development team is acknowledged for its support and efficiency. We would also like to thank LABOGENA for performing SNP genotyping and the Toulouse Midi-Pyrénées bioinformatics platform that partly supported computations. This project was funded by the French Region Centre, the Animal Genetics department of INRA, ANR, and APISGENE (SheepSNPQTL project) and EU grants (FEOGA 024/m/1873) **[AU: location (city, country) needed for grant information]**. G. Sallé holds an INRA grant (Departments of Animal Health and of Animal Genetics).

[2]Corresponding author: Guillaume.Salle@toulouse.inra.fr

Received January 12, 2012.
Accepted June 19, 2012.

the French Romane breed, formerly known as INRA401 (Terefe et al., 2007).

Taking advantage of the resistance contrast between the subtropical Martinik Black Belly breed (**MBB**) and the Romane breed (**RMN**) in a backcross population consisting of 1,275 experimentally infected lambs, the aim of this study was to perform in-depth QTL detection for the resistance to *H. contortus*. Extensive phenotyping was achieved through measurement of fecal egg count (**FEC**), packed cell volumes (**PCV**), total worm burden, worm female length, abomasal pH, blood **[AU: do you want to use "serum" for consistency here? Please advise]** and mucosal specific G immunoglobulins (**IgG**), and serum pepsinogen. Genotyping strategy firstly was a microsatellite selective genotyping of a subset of 332 lambs with 160 microsatellite markers before switching to SNP DNA markers from the Illumina Ovine SNP50 BeadChip genotyped for 1,000 lambs.

MATERIALS AND METHODS

All animals were kept indoors, handled with care, and managed as a commercial flock following the INRA ethics policy. At the end of the experimental infection, animals were slaughtered at the INRA-Nouzilly abattoir following the EU rules.

Study Populations

Backcross population. A backcross design between resistant Martinik Black Belly and susceptible Romane breeds was set up as follows (Table 1): 5 F_1 (MBB × RMN) rams were backcrossed by intrauterine AI with 600 purebred RMN ewes at the INRA experimental farm, La Sapinière (UE0332, Osmoy, France). These inseminations resulted in the birth of 1,046 male and female lambs forming a first population born in 2003 and hereafter denoted **BC1**. According to the preliminary QTL mapping study performed in the BC1 population, 4 out of the 5 original F_1 rams were identified as segregating for most of the QTL. In 2006, these 4 sires were mated to 134 RMN ewes that gave birth to 229 backcross lambs (denoted **BC2**).

Additional Purebred Populations. To estimate the linkage disequilibrium (**LD**) extent in the 2 pure breeds that composed the backcross population, additional animals were used. A 90-individual MBB population structured in nuclear families (i.e., sire, dam, and progeny) was selected to be representative of the entire population. The RMN population was composed of an 8-family granddaughter design of 1,050 individuals.

Table 1. Summary of experimental design for each family available[1,2]

	Total no. lambs Infected	No. lambs with extreme FEC_a (used for microsatellite selective genotyping)	Available SNP data after quality control
Sire No.			
14952	268 (48)	72	249
14971	191 (0)	64	0
14976	301 (63)	66	284
14988	269 (75)	64	247
16754	245 (43)	66	220
Total	1,275 (229)	332	1,000
Measured traits[2]	FEC, PCV Peps (for BC2 only)	WB, L, IgGm, IgGs SexR, pH	

[1]For each founder sire, the respective number of progeny, among which the number of extreme resistant/susceptible BC1 (first backcross between resistant Martinik Black Belly and susceptible Romane breeds) progeny used for fine phenotyping, are reported. The respective numbers of BC2 (second Martinik Black Belly × Romane backcross, with rams selected for specific QTL) progeny are indicated in parentheses. Every trait that was measured for each subset of lambs is indicated and number of individuals that were genotyped for the SNP chip (4 families).

[2]FEC, fecal egg count; FEC_a, animal solution of a mixed model equation with the infection rank added to other fixed effects and animal fitted as a random variable; PCV, packed-cell volume; Peps, pepsinogen concentration; WB, worm burden; L, female worm length; IgGm, G immunoglobulins concentration in abomasal mucus; IgGs, G immunoglobulins concentration in serum; SexR, sex ratio in adult worm population; pH, abomasal pH.

Measured Traits

Infection Procedure. Lambs were challenged artificially twice at 3 and 5 mo of age with 10,000 infective *H. contortus* larvae. At the end of the first infection (i.e., 41 d postchallenge; **dpc**), lambs were drenched (7.5 mg of levamisole 5%/kg BW; Virbac S. A., Carros, France) and remained uninfected for an 8-d washout period before being reinfected with the same number of larvae.

FEC and PCV Traits Measured on BC1 and BC2 Populations. Fecal samples were taken twice at 25 and 35 dpc for the 1,275 backcross lambs of the BC1 and BC2 populations. Fecal egg counts were determined following the McMaster technique modified by Raynaud (1970), and FEC values were averaged for each infection (hereafter denoted **FEC12** for the first infection and **FEC34** for the second infection). Blood samples were taken in EDTA-coated tubes to determine PCV at d 0 (denoted **PCV0**). A second PCV measure was done at 41 dpc in both infections in BC1 population and at 35 dpc in BC2 population (denoted **PCV1** and **PCV2** for 1st and 2nd infections, respectively).

Fine Phenotyping of 332 BC1 Lambs with Extreme FEC. After the second infection, the 15% most resistant and 15% most susceptible BC1 lambs of each ram family were selected according to their overall FEC measurements. These selected subsets of 332 lambs

were slaughtered on d 42 ± 2 after the second infection to gather additional phenotypes. Measurement techniques have been fully described elsewhere (Lacroux et al., 2006). Total worm burden (**WB**) was determined and the sex ratio (**SexR**) of adult worms (females:males) was computed. Total lengths (**L**) of 20 randomly chosen female worms per lamb were measured using a video camera. The abomasal pH (**pH**) was evaluated by using pH paper in direct contact with the mucous layer. Serum and mucus IgG (**IgGs** and **IgGm**, respectively) concentrations were determined by indirect ELISA (Lacroux et al., 2006) performed on serum and abomasal fundic mucosa collected from each animal just after death.

Pepsinogen Concentration Measured in the 229 BC2 Lambs. Serum pepsinogen concentrations were determined using a micromethod for routine determination adapted from Dorny and Vercruysse (1998). This trait is a direct indicator of the mucosal damage caused by *H. contortus* infection. For each infection, the pepsinogen concentration was measured twice at d 0 and at 15 dpc. These measurements are hereafter denoted **Peps1** and **Peps2** for the 1st infection and **Peps3** and **Peps4** for the 2nd infection.

Molecular Quality Checks and Map Construction

DNA extraction and genotyping were performed at LABOGENA (www.labogena.fr; verified November 9, 2012).

Microsatellite Genotypes. A panel of 160 microsatellite markers distributed across the whole genome was used. The 2 groups of BC1 lambs with extreme FEC were genotyped, resulting in a so-called selective genotyping procedure (Lander and Botstein, 1989). In addition to these 332 lambs, their 5 sires were also genotyped (Table 1). Raw microsatellite data were analyzed with Genetic Profiler v. 1.5 software (Amersham Biosciences, Uppsala, Sweden). The relative positions of markers were established with the Cri-map software (Green et al., 1990) using the international sheep map (http://www.ncbi.nlm.nih.gov/mapview; verified November 9, 2012) as well as both the human and bovine genome sequences for a subset of unpublished markers developed by the INRA institute.

SNP Genotypes. The 4 backcross families segregating for most of the QTL found with microsatellites were genotyped with the IlluminaOvineSNP50 Beadchip (Illumina, Inc., San Diego, CA). In total, 1,044 backcross lambs were genotyped, as well as their 4 respective F_1 sires and the 4 MBB founders (Table 1). In addition, the 2 additional purebred MBB and RMN populations were genotyped, hence providing full 50K SNP chip data for 90 MBB and 1,050 RMN individuals.

Individuals with a call rate below 98% were discarded, and a 99.9% technical reliability was established by duplicated genotyping in 50 animals. In addition, Mendelian inconsistencies (i.e., no allele shared in common between a progeny and its sire for a given SNP) were checked. Intrinsic SNP quality criteria were also considered. Useless SNP that had been eliminated in the frame of the sheep HAPMAP project (i.e., SNP that Illumina annotated as abnormal; SNP with minor allele frequencies equal to 0; SNP that displayed discordant genotypes between experiments; or SNP showing Mendelian inconsistencies within the International Mapping Flock) were discarded (J. W. Kijas, personal communication). Within our population, SNP with a call rate < 97% were removed. A minor allele frequency < 1% was applied. A test for Hardy-Weinberg disequilibrium was also considered in purebred populations to eliminate SNP with abnormal behavior ($P < 10^{-6}$). The SNP discarded in purebred populations were also eliminated from the backcross genotype data. Furthermore, SNP for which more heterozygotes than expected (i.e., 50% for a heterozygous sire) were counted, or SNP with too great a recombination rate were not included in the QTL analysis. Sex chromosomes were not considered for analysis. Finally, 1,000 backcross individuals, 939 RMN individuals, 90 MBB sheep, and 42,469 autosomous markers were retained for subsequent analysis.

The SNP positions were obtained from the sheep Genome Browser v. 2.0 on http://www.livestockgenomics.csiro.au/cgi-bin/gbrowse/oarv2.0/ (verified November 9, 2012; Archibald et al., 2010). For mapping purposes, 1 Mbp was considered as equivalent to 1 cM.

Statistical Handling of Phenotypes

Transformations Applied to Phenotypes. Basic statistics and correlations of phenotypes were computed and data were tested for normality using the Shapiro-Wilk test (SAS 9.1.3 Help and Documentation; SAS Institute Inc., Cary, NC). A quasinormal distribution was observed for PCV, L, and IgGs traits. A fourth root transformation was applied to FEC traits (**FEC12t** and **FEC34t** for 1st and 2nd infections, respectively) as well as to pH and IgGm (denoted **pHt** and **IgGmt**). Square root transformation corrected departures from normality for worm burden and sex ratio (**WBt** and **SexRt**, respectively). Finally, the difference between pepsinogen concentrations under a naive state and during infection was considered, and a fourth root was applied (denoted **Peps12t** and **Peps34t** for the 1st and 2nd infections, respectively).

Correction for Fixed Effects. Transformed phenotypes were subsequently corrected for the usually encountered environmental effects (i.e., sex, management group, litter size, and age at infection). The FEC and PCV were also considered as longitudinal traits over both infections by estimating the animal solution of a mixed model equation (SAS PROC MIXED), with the infection rank

added to other fixed effects and animal fitted as a random variable. These traits were denoted **FEC_a** and **PCV_a**. To account for within-animal physiological variation in PCV, PCV values obtained after infection were corrected with PCV0 fitted as a covariable and denoted **PCV1c** and **PCV2c**. Determination of the significant fixed effects was performed using SAS PROC GLM.

Phenotypic Correlations. Phenotypic correlations between transformed traits corrected for environmental effects were calculated using the Pearson correlation coefficient (SAS PROC CORR). In the BC1 flock, immunopathological traits were measured in 2 pools of animals selected on their FEC; this resulted in a bias in the observed correlations between other traits. Therefore, correlation coefficients were estimated by taking into account the increase in variance of the selected population (see formulae in APPENDIX 1).

Methods for Analysis of Marker Data

Comparison of LD and LD Phase. The r^2 LD measure (Hill and Robertson, 1968) was computed within pure breeds between SNP pairs less than 1 Mbp apart using SNP data and 2 specific pure breed populations. The correlation of *r* values across breeds expressed as a function of genomic distance was considered to investigate the persistence of LD phase between the 2o **[AU: 2? or 20? Please fix number]** breeds, as reported by de Roos et al. (2008).

Linkage Analysis. Linkage analysis (**LA**) was performed with microsatellite markers in the selective genotyping design (denoted **LA-micro**) and with the 4-family SNP genotyped dataset (denoted **LA-SNP**). Within each family, the presence of a QTL was tested against the null hypothesis, which was absence of a QTL at every 0.1-cM interval (approximately 1 Mbp/1 cM), by likelihood computation using the QTLMAP software (Elsen et al., 1999). Chromosome-wise significance was determined for each chromosome-trait pair by testing with 10,000 permutations (Churchill and Doerge, 1994). To prevent any overestimations of thresholds due to the selective phenotyping occurring for some traits, permutations were restricted to animals that were both genotyped and phenotyped for the considered traits. These permutations were used to calculate the genome-wise suggestive threshold (1 false-positive result was expected for a whole genome scan) and genome-wise significant threshold (5% genome-wise significance threshold) as indicated in Lander and Kruglyak (1995). These genome-wise thresholds were obtained by applying the Bonferroni correction $P_{\text{genome-wise}} = 1 - (1 - P_{\text{chromosome-wise}})^{n}$, where n is the number of chromosomes (i.e., 26 in sheep; Knott et al., 1998). Confidence intervals were determined using the 2-LOD (i.e., logarithm of the odds) drop-off criterion (Ooijen, 1992) and assuming 1 LOD = 4.61 LRT (i.e., long terminal repeats; Lynch and Walsh, 1998). Estimated QTL effects were corrected for the bias due to selective genotyping (see formulae in APPENDIX 2) following the recommendations of Bovenhuis and Spelman (2000). The QTL effects were hence expressed in phenotypic SD (σ_p) corrected for selective phenotyping as described in APPENDIX 1. A *t*-test was performed to identify the 5% significant QTL effects within each sire family.

Association Analysis and Joint Linkage and Association Analysis. The LD-decay model proposed by Legarra and Fernando (2009) was implemented in the QTLMAP software and applied to our data. In this application of the LD-decay model, observed phenotypes were regressed on the conditional probability of having inherited a 4-SNP haplotype from each of its 2 parents, the substitution effect of the sire haplotypes being weighted by the transmission probabilities. In addition, a joint association and linkage analysis (**LDLA**) was performed to take advantage of both LD and pedigree information in the experimental population (Legarra and Fernando, 2009). In this LDLA model, within sire QTL effects are added to the sire haplotypes effects of the LD-decay model to account for a possible between sire variability of the QTL effect beyond that reflected by the haplotype.

For both genome-wide association studies (**GWAS**) and LDLA analyses, the minor haplotype frequency was set at 1% for analysis of the FEC and PCV traits and at 5% for other traits because fewer animals were measured, hence ensuring at least 10 observations for estimation of the maternal haplotype effect.

In both models described by Legarra and Fernando (2009), it is assumed that every founder originated from the same common base population. Because our working population was a mixture of 2 breeds, additional analysis was performed by clustering founders haplotypes according to their breed origin (i.e., 2 identical haplotypes were considered to be different if originating from different breeds). This is similar to Pérez-Enciso and Varona (2000). These analyses were annotated with a *b* index (**GWASb** and **LDLAb** for association and joint analysis, respectively).

The chromosome-wise P values were estimated for each trait assuming the LRT statistics asymptotically followed a c^2-distribution with *k* degrees of freedom, *k* being the number of QTL effects (Piepho, 2001). Hence, *k* was equal to the number of haplotypes minus 1 for GWAS and the number of haplotypes plus the number of families minus 1 for LDLA. Genome-wise *P* values were derived by applying a Bonferroni correction as described above. Any association reaching the 1% genome-wise significant *P* value was declared significant. Associations reaching the 5% genome-wise significant

threshold were only reported if a suggestive QTL had already been found for the same trait by LA-SNP analysis.

RESULTS

Phenotypes and Correlations

A summary of PCV and FEC statistics for challenged lambs is reported in Table 2. During the primary infection, large FEC values were measured and the mean PCV of infected lambs decreased by 11%. Upon reinfection, FEC were significantly reduced in comparison with the 1st infection (–76% on average) and mean PCV were less reduced (Table 2). After worm development in the abomasum, mean pepsinogen concentrations followed a significant 2.6- and 3-fold increase between d 0 and 15 dpc in 1st and 2nd infections, respectively.

At the end of the second infection period of the BC1 population, a pool of 332 BC1 lambs with extremely large (susceptible; **S group**) or small (resistant; **R group**) FEC values ($P < 10^{-4}$) were assessed for additional phenotypes (i.e., characteristics of worm populations, abomasal pH, and IgGs and IgGm concentrations; Table 3). The mean total number of worms in the S group was 3.6-fold greater than in the R group. The R group also exhibited a greater number of immature stages. Moreover, female worms measured in the S lambs were significantly longer ($P < 10^{-4}$). SexR in adult worms were slightly different between the 2 groups with more females counted in the R group ($P = 0.02$). Resistant lambs maintained greater PCV values during both infections ($P < 10^{-4}$) but exhibited lesser serum IgG concentrations ($P < 2.10^{-4}$). No significant differences were observed between R and S groups as regards to abomasal pH and mucosal IgG values.

Phenotypic correlations between the different variables in the BC1 population are listed in Table 4. Mean FEC and PCV values were negatively correlated in both infections whatever the population, whereas PCV2 was positively correlated to PCV1. Total worm burden, length of female worms, and serum IgG concentrations were all positively correlated to FEC and negatively correlated to PCV in both infections. In addition, it was observed that the more worms counted in the abomasum, the greater the proportion of males, the longer the females, and the greater the serum IgG concentration. Most of the other correlations were not significant. Regarding pepsinogen concentrations measured in BC2 lambs, 2 significant correlations of –0.21 and –0.24 were found between FEC34t and Peps4 and PCV1 and Peps3, respectively.

Table 2. Basic statistics of packed-cell volume (PCV) and fecal egg counts (FEC) for the whole backcross population (n = 1,275 lambs) and summary statistics of pepsinogen concentration for the BC2 flock (n = 229 lambs), before and after 2 successive infections with *Haemonchus contortus*[1]

Trait	n	Mean	±SD	Minimum	Maximum
FEC and PCV in the whole population					
Before challenge					
Packed cell volume, %	1,245	39	5	26	55
First challenge					
FEC mean of 25 and 35 dpc,[2] eggs/g of feces	1,224	11,034	10,964	0	87,950
Packed cell volume before drenching, %	1,229	33	8	10	55
Second challenge					
FEC mean of 25 and 35 dpc,[2] eggs/g of feces	1,190	2,620	4,322	0	42,667
Packed cell volume before drenching, %	1,191	35	7	6	55
Pepsinogen concentration in the BC2 population					
Before any challenge, U/L	200	0.32	0.19	0	1.51
35 d after the 1st challenge, U/L	200	0.83	0.37	0	2.43
Before the 2nd challenge, U/L	201	0.35	0.16	0	1.03
35 d after the 2nd challenge, U/L	201	1.02	0.56	0.19	3.43

[1]BC2 represents the second of Martinik Black Belly × Romane backcross, with rams selected for specific QTL.

[2]dpc, days postchallenge.

Table 3. Egg excretion and parasitological examination of necropsied animals (n = 332) from the extreme resistant and susceptible groups (R and S, respectively)

	R group (n = 169)		S group (n = 163)		
Trait	Mean	±SD	Mean	±SD	*P* value
Total worm burden	1,141	1,277	4,117	2,088	10^{-4}
L4 larvae, %	8%	18%	4%	10%	10^{-2}
Juveniles, %	6%	13%	2%	4%	10^{-3}
Adult males, %	38%	16%	45%	9%	10^{-4}
Adult females, %	49%	18%	49%	7%	NS
Sex ratio in adult worms	1.39	0.83	1.19	0.81	2.10^{-2}
Adult female length, µm	17,080	2,158	18,548	1,337	10^{-4}
pH in abomasum[1]	3.5	0.8	3.3	0.9	NS
IgG in serum[2]	44	26	55	28	2.10^{-4}
IgG in mucus[2]	53	71	56	76	NS
PCV0, %	40	5	40	6	NS
PCV1, %	37	7	32	9	10^{-4}
PCV2, %	39	7	32	8	10^{-4}
FEC12,[3] eggs/g of feces	4,427	6,616	21,463	15,572	10^{-4}
FEC34,[3] eggs/g of feces	292	709	7,014	6,705	10^{-4}

[1]For pH, only 119 resistant and 115 susceptible lambs were measured.

[2]IgG (G immunoglobulins) concentrations are expressed as a percentage of the basal value.

[3]FEC12, fecal egg count mean between 25 and 35 d after the 1st challenge; FEC34, fecal egg count mean between 25 and 35 d after the 2nd challenge.

Table 4. Phenotypic correlation (Pearson coefficient) between FEC, PCV, and parasitological examination traits[1]

	FEC12	FEC34	PCV1	PCV2	FEC_a	PCV_a	WB	SexR	L	pH	IgGm	IgGs
PCV0	–0.07*	0.04	0.00	0.01	–0.02	0.06	–0.04	0.04	0.11*	0.00	0.02	–0.03
FEC12		0.28**	–0.33**	–0.14**	0.80**	–0.31**	0.40**	–0.11*	0.26**	–0.11*	0.11*	0.23**
FEC34			–0.16**	–0.42**	0.81**	–0.41**	0.69**	–0.19**	0.33**	–0.12*	–0.05	0.13*
PCV1				0.13*	–0.30**	0.72**	–0.21**	0.02	–0.16*	0.11	–0.10*	–0.18*
PCV2					–0.36**	0.77**	–0.51**	0.15*	–0.14*	0.06	–0.02	–0.19*
FEC_a						–0.45**	0.60**	–0.16**	0.31**	–0.13*	0.02	0.18**
PCV_a							–0.45**	0.09	–0.18*	0.10	–0.07	–0.22**
WB								–0.30**	0.23**	–0.17*	0.03	0.18*
SexR									0.00	0.09	–0.09	–0.02
L										–0.04	–0.07	0.03
pH											0.19*	0.01
IgGm												0.29**

*Significantly different from zero, $P < 0.01$. **[AU: confirm changeup of this footnote]**

Significantly different from zero, $P < 0.0001$. **[AU: confirm changeup of this footnote]

[1]FEC, fecal egg count, 12 and 34 indicate FEC after 1st and 2nd challenge, respectively; PCV, packed-cell volume, 1 and 2 indicate PCV after 1st and 2nd challenge, respectively; FEC_a and PCV_a, within-animal physiological variation accounted for in the analysis; WB, worm burden; SexR, sex ratio in adult worm population; L, female worm length; pH, abomasal pH; IgGm, G immunoglobulins in abomasal mucus; IgGs, G immunoglobulins in serum; [b] indicates that haplotype breed origin was accounted for in the analysis **[AU: I don't see a [b] in the table?]**.

Linkage Analysis with Microsatellites or SNP Markers

Full results are provided for the linkage analysis with microsatellites (Supplementary Data 1; see online version of article to access supplemental material, at http://journalofanimalscience.org) or SNP markers (Supplementary Data 2), as well as a comparison of the 2 types of molecular information (Supplementary Data 5 **[AU: cited out of order, advise on the fix? Also, Supplementary Data 4 is not cited at all]**).

Five-Family Selective Genotyping Design with Microsatellites. Linkage analysis was performed on the animals genotyped with the microsatellite panel (i.e., extreme R and S groups from the BC1 flock). A total of 8 suggestive and 3 significant QTL regions for resistance to *H. contortus* were detected (Table 5, Figure 1A, 1B). The QTL mostly affected FEC traits, PCV, and total worm burden, but also specific variables such as IgGmt, SexRt, L, and pH. Chromosomes 3, 12 (see Figure 1A, 1B), and 23 were tagged as key players as they each carried significant QTL for several correlated traits (FEC12t, FEC34t, FECt_a, or WBt). In addition, these QTL were found in the same families and showed overlapping confidence intervals. Other QTL affecting pH, SexRt, L, and IgGmt were mapped in clearly distinct regions on chromosomes 2, 5, 12, 13, 18, and 19.

Linkage Analysis with SNP Data. Four out of the 5 original families were genotyped with SNP (Table 1). In total, 9 significant QTL were mapped on **[AU: define OAR again here]** OAR5, 7, 10, 12, 13, and 21. As in the selective genotyping design with microsatellites, most of the QTL were related to FEC traits (16 out of the 38 at least suggestive QTL). Mean QTL effects for FEC and PCV ranged between 0.13 and 0.3σ_p, whereas QTL effects estimated on other parasitological phenotypes were greater (Table 5).

Almost 50% of the QTL were clustered on 3 chromosomes (i.e., OAR5, 7, and 12). Of particular interest was the 10-Mbp region on OAR12 (Figure 2A, 2B) that was significantly associated with each of the FEC traits and exhibited some of the greatest effects for this trait (0.22 σ_p for FECt_a and 0.19 σ_p for the 2 other FEC traits). On OAR5, a suggestive QTL region located at the end of the chromosome was repeatedly associated with 7 correlated traits, but the confidence intervals usually spanned more than half of the chromosome (Figure 3A, 3B). It is interesting to note that 4 QTL were mapped on OAR7. A significant QTL with a narrow confidence interval was obtained for FEC12t but other suggestive QTL on this chromosome were distributed in a rather erratic way, either at the right (FEC34t) or at the left end of the chromosome. Last but not least, a QTL region on OAR21 was significantly associated with Peps34t at 37.8 Mbp (i.e., 3 Mbp from another suggestive QTL related to FECt_a; Figure 4). Other chromosomes showing multiple suggestive QTL (i.e., OAR16, 17, 23, and 25) also exhibited at least 10 Mbp between the peaks and/or had large confidence intervals (Table 5). It is worth noticing that no QTL were detected on OAR3 (but detected with microsatellites), but 1 suggestive region was reported for FEC34t on OAR20.

Taking Advantage of LD in the Experimental Population

LD Estimation in Pure Breeds. The LD between SNP markers was rather small in both breeds. The decrease in LD was slightly less in MBB compared with their RMN counterparts (Supplementary Figure 1A). At the average marker spacing (i.e., 57 kb), r^2 value was 0.14 and 0.13 for MBB and RMN, respectively. As expected, phase persistence between the 2 breeds decreased

Table 5. Summary of linkage analyses (LA) studies with microsatellites or SNP data

		Linkage analyses				
		Selective genotyping with microsatellites		SNP data		Avg QTL Effect[3]
OAR	Trait[1]	Significance	Position, cM [CI][2]	Significance	Position, Mb [CI]	
1	PCV_a	–	–	*	146.5 [21.1–236.6]	0.15
2	PCV1c	–	–	*	244.3 [128.4–248.5]	0.19
2	SexRt	*	52 [9–80]	–	–	–
3	FEC34t	*	158 [141–167]	–	–	0.68
3	WBt	**	155 [144–162]	–	–	0.86
4	L	–	–	*	45.6 [32.6–72.5]	0.36
4	pHt	*	73 [69–86]	–	–	0.89
5	FEC12t	–	–	*	86.7 [35.1–94.2]	0.16
5	FEC_a	–	–	*	86.7 [35.2–94.1]	0.17
5	IgGst	–	–	*	87.3 [60.5–94.2]	0.54
5	WBt	*	82 [72–89]	*	80.6 [37.4–94.4]	0.42
5	PCV1c	–	–	*	68.3 [32.2–91.3]	0.15
5	PCV2c	–	–	*	87.4 [7.6–96.7]	0.14
5	PCV_a	–	–	**	71.8 [53.3–92.7]	0.15
7	FEC12t	–	–	**	46.4 [28.8–52.4]	0.17
7	FEC34t	–	–	*	97.3 [21–100]	0.15
7	FEC_a	–	–	*	46.3 [20.4–50.5]	0.17
7	L	–	–	*	13.8 [4.4–21.7]	0.55
9	FEC_a	–	–	*	52.3 [24.8–94.9]	0.13
10	Peps34t	–	–	**	83.5 [80.3–83.5]	0.66
12	FEC12t	**	55 [17–72]	**	56 [45.8–60.2]	0.19
12	FEC34t	*	59 [53–65]	**	47 [35.6–58.3]	0.19
12	FEC_a	**	58 [52–68]	**	46.3 [36.5–57.4]	0.22
12	PCV2c	–	–	*	45.4 [29.9–57.7]	0.20
12	PCV_a	–	–	*	35.6 [19.7–56.8]	0.30
12	pHt	*	77 [72–82]	–	–	0.84
13	FEC34t	–	–	**	72.3 [70.1–77.8]	0.22
13	IgGmt	*	109 [98–117]	*	0.6 [0.2–81.6]	0.30
14	Peps12t	–	–	*	61.2 [54.3–62.2]	0.36
15	FEC12t	no marker	no marker	*	43.5 [1.6–52.3]	0.18
16	FEC34t	–	–	*	27.2 [10.8–41.1]	0.18
16	WBt	–	–	*	19.2 [11.2–29.1]	0.44
17	pHt	–	–	*	63.7 [47.5–64.7]	0.31
17	PCV1c	–	–	*	18.1 [12.6–66.7]	0.13
17	PCV_a	–	–	*	65.1 [2.8–72.4]	0.17
18	L	*	43 [49–81]	–	–	0.53
19	L	*	51 [61–66]	–	–	0.48
20	FEC34t	–	–	*	31.3 [19.8–41.5]	0.17
21	FEC_a	–	–	*	41 [1.6–46.8]	0.14
21	Peps34t	–	–	**	37.8 [31.8–46.1]	0.79
23	FEC34t	*	51 [40–80]	*	59.6 [0.8–62.7]	0.15
23	FEC_a	*	54 [43–74]	*	32.2 [0.3–62.7]	0.27
23	WBt	*	63 [52–70]	*	44.1 [15–59.8]	0.58
25	SexRt	–	–	*	39.5 [0.4–40.7]	0.56
25	PCV2c	–	–	**	41.4 [16.6–44]	0.17

*Suggestive threshold.

**5% genome-wise significant.

[1]For FEC traits, results for SNP data are based on the four families backcross population. FEC, fecal egg count, 12 and 34 indicate FEC after 1st and 2nd challenge, respectively; PCV, packed-cell volume, 1 and 2 indicate PCV after 1st and 2nd challenge, respectively, and c indicates values corrected with PCV0 fitted as a covariable; WB, worm burden; SexR, sex ratio in adult worm population; L, female worm length; pH, abomasal pH; IgGm, G immunoglobulins in abomasal mucus; IgGs, G immunoglobulins in serum; Peps, pepsinogen; t, fourth root transformation of the variable; _a, within-animal physiological variation accounted for.

[2]CI, confidence interval.

[3]From the SNP analysis if the QTL were found in both analyses; given in phenotypic SD.

rapidly with genomic distance: *r* coefficients decreased from 90% at 10 Kb to 50% at the average SNP spacing, and a 50% correlation was observed at the average SNP spacing (Supplementary Figure 1B). The LD phase persistence between MBB and RMN breeds decreased rapidly with genomic distances, and even at relatively short distances (<10 Kb) some correlations between *r* exhibited opposite signs (Supplementary Figure 1C).

Association Analysis and Joint Linkage and Association Analysis. Both association and joint linkage and association analyses were performed twice, either considering or not the breed origin of the haplotypes. Significant QTL are reported in Table 6, and Manhattan plots of the LDLA analysis for FEC12t and FEC34t are presented in Figures 5 and 6. Extensive results of the LD-based analyses performed can be found in Supplementary Data 3 (GWAS and LDLA analyses) and 4 (GWASb and LDLAb analyses).

In total, 41 haplotype-trait associations were declared significant, 24 of which were detected by LDLA analysis only (Table 6). By contrast, GWAS detected only 1 significant QTL, whereas both GWASb and LDLAb performed similarly (n = 11 and 15 significant QTL, respectively).

The chromosomes already highlighted by LA analysis (i.e., OAR5, 7, 12, 13, and 21) showed consistent significant associations using 3 different analysis methods (Figures 2 to 6 and Table 6). In addition, at least 4 trait-haplotype associations were found on OAR5, 12, and 21 (Figures 2, 3, and 4). An association hot-spot was associated with FEC, PCV, and pepsinogen concentration at the end of OAR5, and OAR12 was the only chromosome to be associated with both FEC traits (Table 6). The QTL region found by LA for Peps34t on OAR21 was associated with both Peps12t and Peps34t and pinpointed a 1-Mbp interval between 36.7 and 37.7 Mbp (Tables 5 and 6). Three additional associations were found on OAR20 and OAR23, but the SNP haplotypes were at least 10 Mbp apart from each other. Other chromosomes (OAR1, 2, 4, 8, 14, 17, 18, 22, 25, and 26) also exhibited 1% genome-wise significant associations but with the LDLA method only (Table 6). Among these, no previous QTL had been mapped using LA on chromosomes 8, 22, and 26 hence suggesting they segregate in the RMN breed only or that they are spurious signals (Table 5).

DISCUSSION

Overview

The aim of our study was to provide an exhaustive characterization of the genetic basis of breed differences in the response to *H. contortus*. This was not only achieved by using dense molecular information but also by studying a broad range of phenotypes (i.e., FEC, PCV, worm burden, length of females, IgG, and pepsinogen concentration) in more than 200 animals. After a first QTL mapping exercise with a selective genotyping strategy using microsatellites, 50k SNP markers were genotyped for 4 backcross families (1,000 genotyped individuals after quality control) in an attempt to more precisely map the QTL locations. This work complements another study by Kemper et al. (2011) that also used ovine SNP data to investigate the genetic architecture of host resistance to *Trichostrongylus colubriformis* and *H. contortus* in a large mixed breed population (more than 2,000 animals), again using FEC data.

Infection Dynamics

Considering basic traits, a sharp decrease in egg counts was observed upon reinfection, whereas large FEC values were associated with low PCV. In addition, the dynamics and intensity of FEC were similar between the 2 BC1 (mean FEC12 = 11,697 and mean FEC34 = 2,706) and BC2 flocks (mean FEC12 = 8,665 and mean FEC34 = 2,347). However, no increase in FEC was reported between 28 and 35 dpc during the first exposure period to *H. contortus* in the BC1 flock. It may be possible that the onset of the immune response in BC2 lambs at the 1st infection was not prompt enough to regulate worm growth. The R subgroup exhibited a smaller total worm burden and more immature worms in comparison with the S group. In addition, the length of female worms, which is correlated to female worm fertility (Stear et al., 1999), was shorter in the R group. Hence, the resistance of backcross lambs relied both on a reduction in worm burden and female fertility. Interestingly, this is consistent with findings in Scottish Blackface, in which the within breed variation of the control of another abomasal trichostrongylid (*Teladorsagia circumcincta*) is due to both a reduction in female fertility and worm burden (Stear et al., 1996). By contrast, Good et al. (2006) reported that the differences observed between Texel and Suffolk were only due to a lessening of the worm population. Interestingly, the control of *H. contortus* in the backcross population, as well as the differences observed between RMN and MBB breeds (Terefe et al., 2007), appear to follow the same pattern as reported within Scottish Blackface. The IgG titers were greater in the S group and positively correlated with FEC, hence strongly contrasting with results obtained in RMN lambs experimentally infected with *H. contortus* (Lacroux et al., 2006). This discrepancy may lie in the smaller number of animals (n = 26) considered for correlation computation in the study performed by Lacroux et al. (2006).

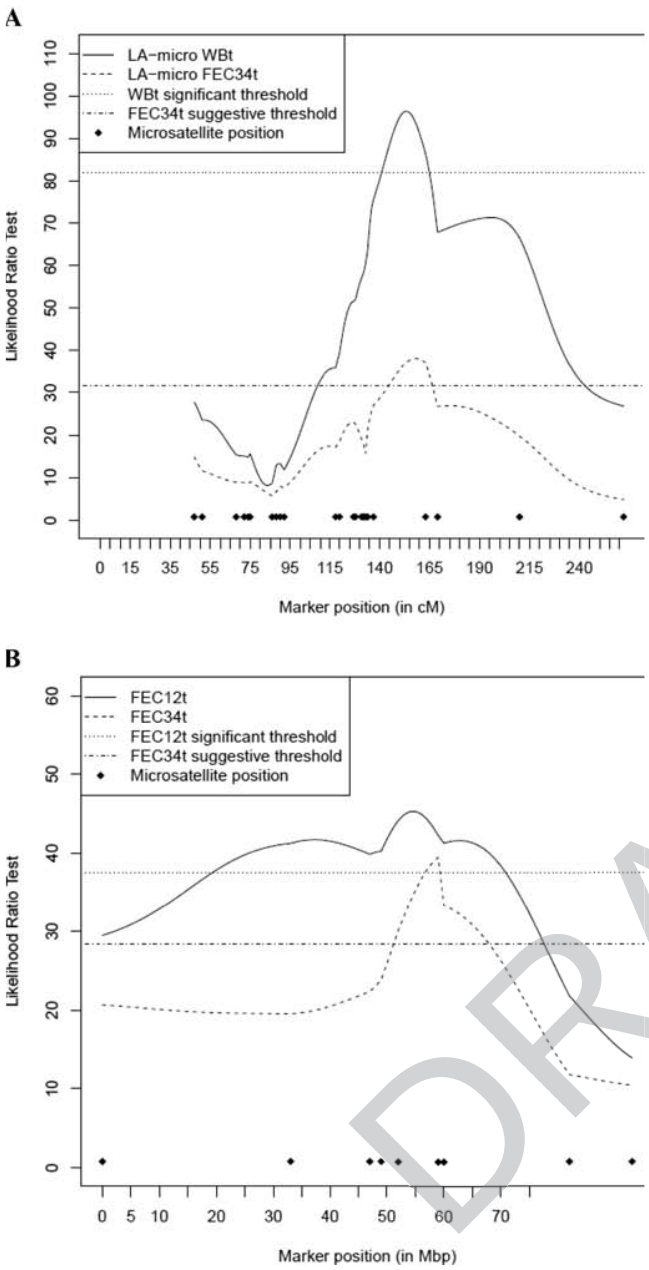


Figure 1. Likelihood profiles of linkage analyses on (A) OAR3 for worm burden (WBt) and fecal egg count at 2nd infection (FEC34t) and on (B) OAR12 for fecal egg count at first and second infection. Marker positions are indicated with a diamond symbol, and estimated thresholds are indicated with horizontal dotted and dashed-dotted lines. In Panel A, the likelihood ratio test statistics of a linkage analysis performed on fecal egg count at 2nd infection (LA-micro FEC34t, dashed line) and worm burden (LA-micro WBt, solid line) measured on the 332 animals from the selective genotyping design is plotted against microsatellites positions (in cM) along OAR3. In Panel B, the likelihood ratio test statistics of linkage analyses performed on fecal egg count at 1st infection (solid line) and 2nd infection (dashed line) measured on the 332 animals from the selective genotyping design is plotted against microsatellites positions (in cM) along OAR12.

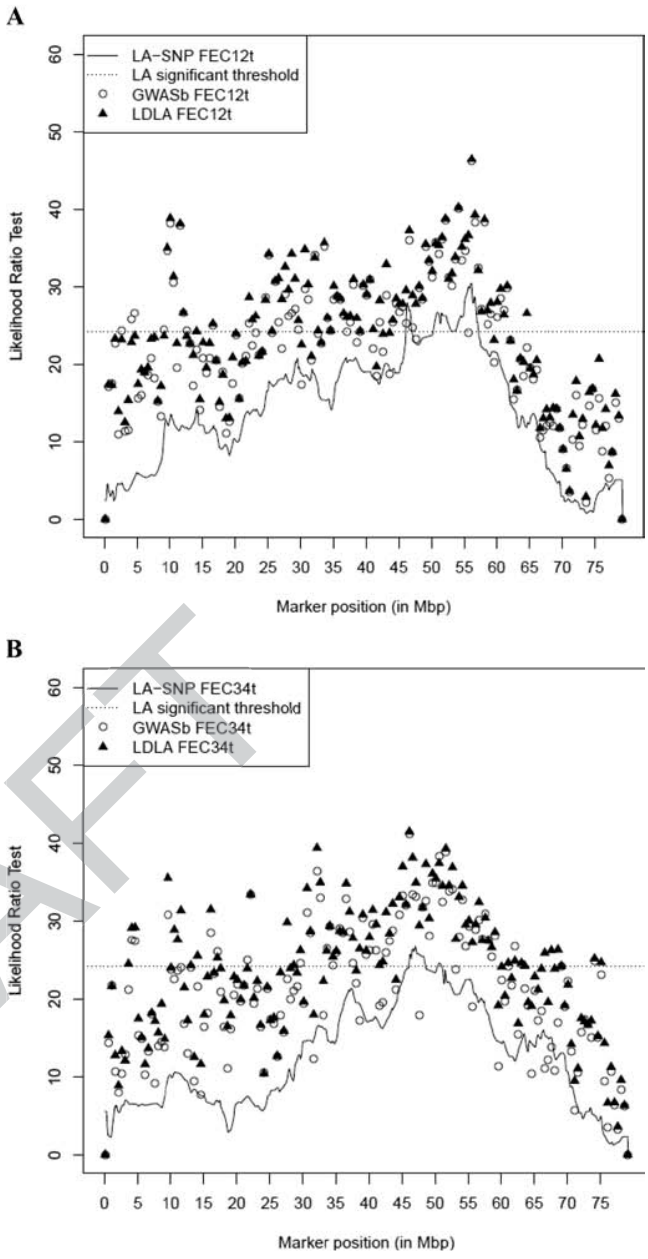


Figure 2. Comparison of the likelihood ratio test profiles obtained on OAR12 for fecal egg count at (A) first and (B) second infection with the linkage, association, or joint linkage and association analyses with SNP data. In Panel A, the likelihood ratio test statistics of linkage (solid line), association (circles), and joint linkage and association analyses (triangles) that a QTL affects fecal egg count at second **[AU: first?]** infection in the SNP-genotyped backcross population is plotted along OAR12. The estimated suggestive threshold of the linkage analysis is indicated by a dotted horizontal line. In Panel B, the likelihood ratio test statistics of linkage (solid line), association (circles), and joint linkage and association analyses (triangles) that a QTL affects fecal egg count at second infection in the SNP-genotyped backcross population is plotted along OAR12. The estimated suggestive threshold of the linkage analysis is indicated by a dotted horizontal line.

An Extensive Genome Scan for Resistance to H. contortus

In total, LA mapped 23 different QTL regions on 20 chromosomes, and LD-based analysis found 1% genome-wise significant QTL on 22 chromosomes. Results from linkage analysis are in good agreement with other QTL mapping studies in which relatively few significant QTL were detected (Crawford et al., 2006; Davies et al., 2006; Gutierrez-Gil et al., 2009; Dominik et

Table 6. QTL detected in at least one linkage disequilibrium (LD)-based analysis at the 5% genome-wise significance (GWS) threshold

OAR	Trait[1]	Flanking SNP[2]	Pos.[3]	P_{GWAS}[3,4]	P_{GWASb}	P_{LDLAa}	P_{LDLAb}
1	PCV1c	OAR1_45014772.1 – OAR1_45507542.1	43.6			*	
2	Peps1	OAR2_10601656.1 – OAR2_10799411.1	11.7			**	**
4	L	OAR4_37940876.1 – OAR4_38205790.1	35.2			**	*
4	pHt	OAR4_86339728.1 – OAR4_86628139.1	80.7			**	
5	FEC12t	OAR5_92638062.1 – OAR5_92975517.1	85.1		*	**	
5	IgGst	OAR5_67605574.1 – OAR5_67883800_X.1	62.1			*	
		OAR5_100699982.1 – DU183841_402.1	93.1		*		**
5	PCV2c	OAR5_98137778.1 – OAR5_98330992.1	90.6			**	
5	PCVt_a	OAR5_94568486.1 – OAR5_94826280_X.1	87.1		**	**	**
		s36267.1 – OAR5_96703012.1	89.1	*			
5	Peps1	s06286.1 – s29268.1	94.6			**	
5	Peps2	OAR5_114751184.1 – s38472.1	106.1			**	**
7	FEC12t	OAR7_36947114.1 – OAR7_37429300.1	32.9			**	
7	L	OAR7_15034944.1 – s39389.1	14.9		**	**	**
		s62332.1 – OAR7_17669851.1	16.9			**	**
8	PCVt_a	OAR8_49745739.1 – OAR8_50006699.1	46.5			**	
9	IgGst	OAR9_85325486.1 – s48117.1	81.0			**	
9	Peps1	OAR9_88974726.1 – OAR9_89299027.1	84.5			**	
12	FEC12t	OAR12_62226914.1 – s68186.1	56.1		*	**	*
12	FEC34t	OAR12_51099743.1 – OAR12_51339516.1	46.1	*	*	**	*
12	FECt_a	OAR12_36077499.1 – s41448.1	32.1	*			
		s23035.1 – OAR12_56589339.1	51.1		**	**	**
12	Peps2	OAR12_40508365.1 – OAR12_41603863.1	36.6			**	
13	FEC34t	s05259.1 – s09612.1	70.7		**	**	**
		s05603.1 – s43133.1	72.2	*		*	
14	IgGmt	OAR14_48832510.1 – s30682.1	46.6			**	
14	PCVt_a	OAR14_1357526.1 – s24656.1	1.1			**	
14	Peps1	OAR14_32272439.1 – OAR14_32706322.1	31.1			**	
15	PCV2c	OAR15_40719719.1 – OAR15_40926306.1	39.2			**	**
15	pHt	s66581.1 – s54590.1	65.2			**	
16	FEC34t	s04660.1 – OAR16_26684182.1	24.7			*	
16	WBt	OAR16_16777017.1 – OAR16_17276041.1	15.7			**	
		s33208.1 – s31223.1	16.7		**	*	*
17	Peps2	OAR17_67650184.1 – s10326.1	62.6			**	
20	FEC12t	OAR20_24357620.1 – OAR20_24626067.1	23.2	*		*	*
20	FEC34t	s69570.1 – OAR20_32868803.1	29.2			*	
20	PCVt_a	s04766.1 – OAR20_44320222.1	40.2			**	
21	IgGmt	s27845.1 – OAR21_14592163.1	12.7			**	
21	L	OAR21_22572149.1 – OAR21_23049634.1	20.2			**	
21	Peps1	s54156.1 – s26955.1	36.7			**	
		s44626.1 – s61819.1	37.2		**		
21	Peps2	s44626.1 – s61819.1	37.2	**	*	*	*
		s26955.1 – s39524.1	37.7		**	**	
22	L	OAR22_23894546.1 – OAR22_24105777.1	20.3			**	
23	FEC34t	OAR23_61434545.1 – OAR23_61932991.1	58.3			**	
23	FECt_a	s72843.1 – OAR23_16996616.1	15.8			*	
23	WBt	s13806.1 – s71481.1	48.3	*		**	**
25	PCV2c	OAR25_43636572.1 – s50428.1	40.1			*	
26	IgGmt	OAR26_21857857.1 – OAR26_22456940.1	18.7			**	

*5% genome-wise threshold.

**1% genome-wise threshold.

[1]FEC, fecal egg count, 12 and 34 indicate FEC after 1st and 2nd challenge, respectively; PCV, packed-cell volume, 1 and 2 indicate PCV after 1st and 2nd challenge, respectively, and c indicates values corrected with PCV0 fitted as a covariable; WB, worm burden; SexR, sex ratio in adult worm population; L, female worm length; pH, abomasal pH; IgGm, G immunoglobulins in abomasal mucus; IgGs, G immunoglobulins in serum; Peps, pepsinogen; t, fourth root transformation of the variable; _a, within-animal physiological variation accounted for.

[2]SNP flanking the haplotypes with significant association.

[3]QTL positions, given in Mbp.

[4]GWAS and GWASb stand for association analysis; joint linkage and association analyses are denoted LDLA and LDLAb; the b indicates that haplotypes breed origin was taken into account in the analysis.

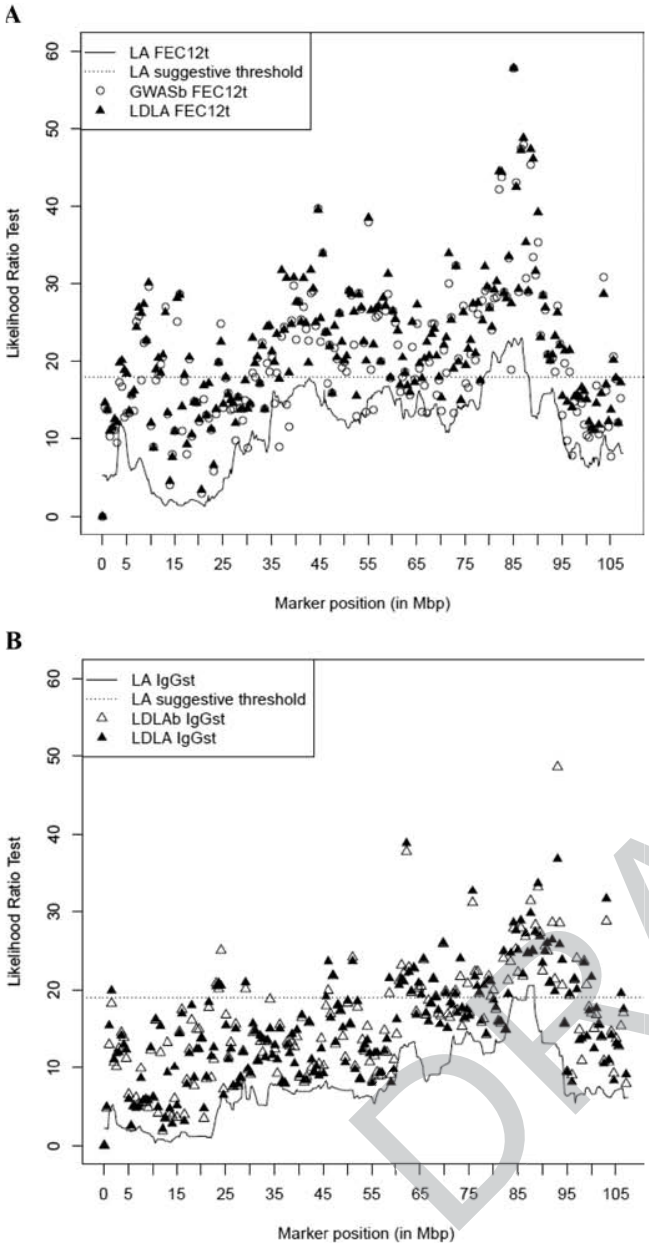


Figure 3. Likelihood ratio test profiles for fecal egg count at 1st infection and the serum G immunoglobulins (IgG) concentration plotted against SNP position on Chromosome 5 for linkage analyses and linkage disequilibrium (LD)-based analyses. In Panel A, the likelihood ratio test statistic that a QTL affects fecal egg count at first infection (FEC12t) is plotted against SNP position (in Mbp) along OAR5. The linkage analysis (solid line), association analysis (circle), and joint linkage and association analysis (filled triangle) are represented. In Panel B, likelihood ratio test profiles of the linkage analysis (solid line) and the joint linkage and association analysis with (LDLAb, filled triangle) or without (LDLA, open triangle) considering the breed origin of the haplotypes are plotted for IgG. The estimated suggestive threshold of the linkage analysis is indicated by a dotted horizontal line.

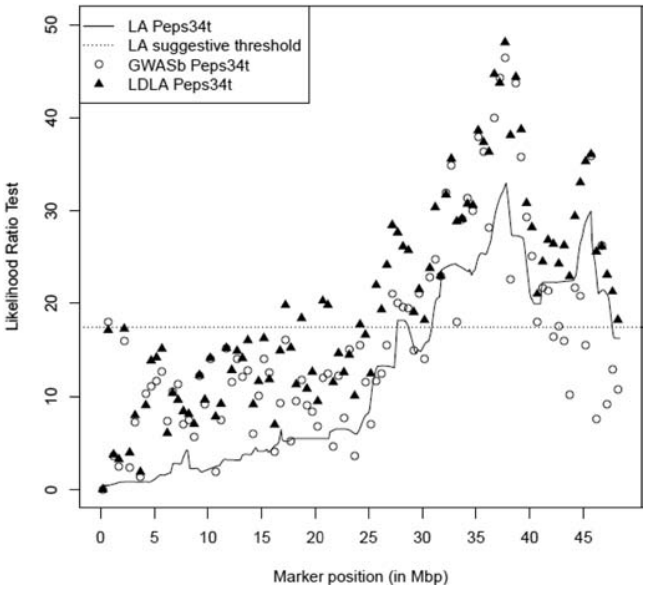


Figure 4. Likelihood ratio test profiles of the linkage analysis with SNP data and both association and joint association and linkage (LDLA) analyses on OAR21 for the pepsinogen concentration measured at 2nd infection. The likelihood ratio test statistic that a QTL affects pepsinogen concentration at 2nd infection is plotted against SNP position (in Mbp) along OAR21 for the linkage analysis (solid line), association analysis (circle), and joint linkage and association analysis (triangle). The estimated suggestive threshold of the linkage analysis is indicated by a dotted horizontal line.

and 65 SNP significantly affecting *T. colubriformis* and *H. contortus* FEC), each exhibiting small to moderate effects (from 0.02% to 0.48% of the phenotypic variation). These effects are in strong contrast to the FEC-associated QTL detected by LA-SNP analysis that explained between 1.7 and 7.5% of the phenotypic variation. By contrast, the QTL effects estimated by LA-SNP analysis were in good agreement with other linkage analyses

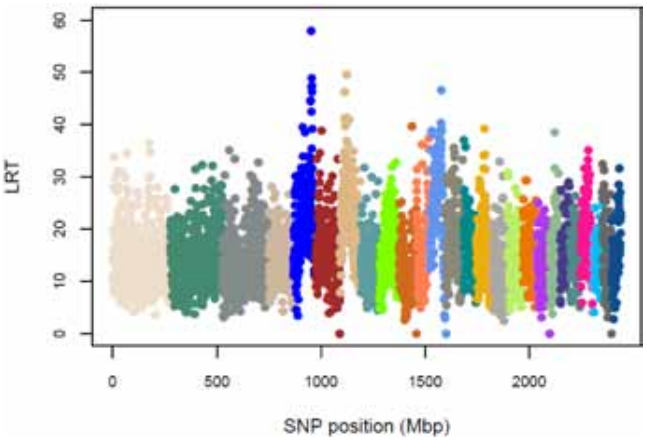


Figure 5. Manhattan plot of the likelihood ratio test values obtained for fecal egg count at 1st infection with the joint association and linkage analysis (LDLA). The likelihood ratio test that a QTL affects FEC at the 1st infection in the backcross population obtained with the LDLA analysis is plotted against the corresponding SNP haplotype position (in Mbp) along the genome. Each plot corresponds to a haplotype of 4 consecutive SNP and each color corresponds to a chromosome from 1 to 26. See online version to view figure in color.

al., 2010), whereas in some studies only suggestive QTL could be detected (Beh et al., 2002; Beraldi et al., 2007). Using ovine SNP data, Kemper et al. (2011) showed that resistance to GIN followed the same infinitesimal model as other complex traits (Hayes and Goddard, 2001; Cole et al., 2009). Indeed they found numerous loci (99 SNP

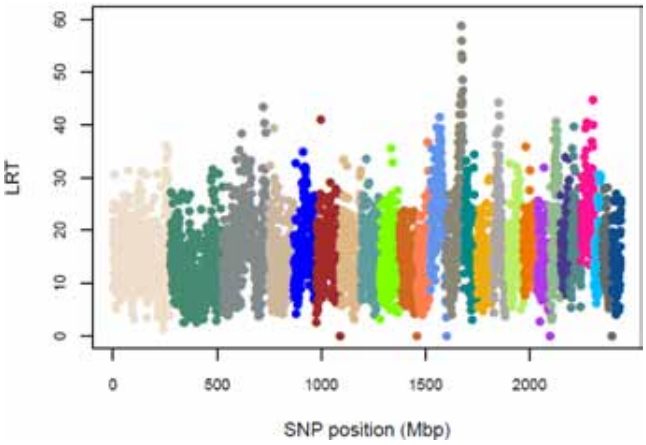


Figure 6. Manhattan plot of the likelihood ratio test values obtained for fecal egg count at 2nd infection with the joint association and linkage analysis (LDLA). The likelihood ratio test that a QTL affects fecal egg count at the 2nd infection in the backcross population obtained with the LDLA analysis is plotted against the corresponding SNP haplotype position (in Mbp) along the genome. Each plot corresponds to a haplotype of 4 consecutive SNP and each color corresponds to a chromosome from 1 to 26. See online version to view figure in color.

(Gutierrez-Gil et al., 2009; Marshall et al., 2009), but somewhat less than the 6 to 13% reported by linkage analysis in other pure breed studies (Davies et al., 2006; Dominik et al., 2010; Matika et al., 2011).

Comparing Linkage Analysis with Microsatellites and SNP

Both the increase in the size of the population studied (from 332 selected extreme BC1 to the 1,000 SNP genotyped BC1 and BC2 individuals) and the denser molecular information led to the detection of 3 times more QTL than in the initial selective genotyping design. The effects of switching from microsatellite to SNP genotypes in terms of power and precision of the linkage analysis was investigated (Supplementary Data 5) as follows. Additional linkage analyses with either microsatellites or SNP genotypes were performed as described in Materials and Methods, using phenotypes of lambs genotyped for both microsatellites and SNP markers (i.e., individuals with extreme FEC_a from 4 families; n = 268). These additional findings showed that the dense genome-wide SNP coverage resulted in the detection of new QTL on chromosomes (i.e., OAR7, 9, 14, 15, and 21) with rather small microsatellite density (on average, 1 microsatellite every 35 cM), . In addition, some of the most consistent QTL regions mapped on OAR5, 12, and 23 with the complete dataset were also found with this reduced dataset whatever the type of markers (Supplementary Data 5). However, the QTL region on OAR3 achieved a *P* value of only 0.073 with SNP data (Supplementary Figure 2; Table 6) but was found suggestive with microsatellites (Figure 1A and Table 5). Other QTL (on OAR4, 6, 8, 17, 18, 19, and 25) were found with microsatellite markers in the reduced population only. In the case of OAR3, the observed discrepancy might therefore originate in the molecular data itself. For the 7 other QTL that were found with microsatellites in the reduced population, the sampling of progeny may have resulted in an overestimation of the QTL effects known as the Beavis effect (Beavis, 1994). This kind of result was also obtained by Melchinger et al. (1998), who detected some QTL in a small plant population that could not be found in a greater population. In linkage analysis, the precision of mapping is mostly driven by the numbers of observed meioses and the QTL effect (Darvasi et al., 1993). In addition, a relatively large microsatellite density had been achieved on OAR5, 12, and 23 (1 microsatellite every 6 to 12 cM), and SNP genotyping was not expected to provide a greatly improved precision. However, the increased marker density provided by SNP markers slightly narrowed down the confidence intervals of the QTL found in common with microsatellites and SNP on OAR5, 12, and 23 (between –13 Mbp to –47.7 Mbp from microsatellites to SNP markers, data not shown).

SNP Unlock the Potential of Existing LD

The LD extent in sheep, estimated with either microsatellites (McRae et al., 2002; Meadows et al., 2008) or SNP markers (Kemper et al., 2011) is small, and the r^2 values estimated in MBB and RMN breeds followed the same pattern. Thus, a dense SNP map was necessary to take advantage of existing LD for mapping purposes. In addition, our population was a mixture of 2 breeds as illustrated by the rapid decrease of LD phase persistence between MBB and RMN breeds with genomic distances (Supplementary Figure 1C). This is consistent with the results of de Roos et al. (2008) in Australian Angus and New Zealand Holstein cows. Moreover, even at short distances (<10 Kb,) some correlations between *r* exhibited opposite signs.

Therefore, 2 types of analyses were performed to take this small LD phase persistence between the 2 pure breeds into account. The GWAS and LDLA did not consider the breed origin of the haplotypes and, hence, focused either on QTL with the same effects in the 2 breeds (both MBB and RMN haplotypes were considered to be the same) or QTL segregating in the RMN breed only (because the frequency in the founders of any eventual MBB-specific haplotypes would be too small to be considered in the analysis). By contrast, GWASb and LDLAb aimed at detecting breed specific QTL in accordance with the hypothesis that a different genetic background was responsible for the resistance of MBB to *H. contortus* infection (Yazwinski et al., 1980; Courtney et al., 1985; Aumont et al., 2003). This hypothesis was confirmed for the LA-SNP QTL located on OAR5, OAR12, and OAR13 because the alleles originating from the MBB breed were linked to smaller FEC (the

difference between the RMN grandmaternal chromosome effect and the MBB grandpaternal chromosome effect being positive). Reinforcing these findings, the GWASb model identified 6 QTL on OAR5, 7, 12, 13, 16, and 21 that could not be detected in the GWAS model. In addition, FEC-related QTL and the QTL for female worm length on OAR7 were related to identical-by-state alleles with opposite effects in each breed (data not shown). However, OAR16 exhibited an opposite pattern and suggested that some part of the RMN genome has a better ability to reduce FEC. This finding corroborates the findings of Terefe et al. (2007), who demonstrated that RMN lambs controlled the parasite more efficiently upon reinfection in comparison with the first infection.

Interestingly, a MBB-specific haplotype mapped at 93.1 Mbp on OAR5 exerted the most favorable effect on IgGst. This QTL was detected by GWASb and LDLAb analyses only, hence illustrating the benefits of taking into account the breed origin of the haplotypes. Indeed, LDLA analysis resulted in a lesser LRT than LDLAb (10 points less) and shifted the maximal LRT value 30 Mbp away at 62.1 Mbp (see Figure 3B). It might be possible that 2 different regions, one carrying an old polymorphism common to both breeds (at 62.1 Mbp) and the other a much more recent mutation (at 93.1 Mbp), may affect IgGst. The older mutation might tend to fixation and, hence, show smaller variations between the possible alleles than the more recent mutation.

Contrasting with this QTL, every other GWASb-specific QTL was also detected by LDLA analysis. The LDLA was expected to have more power than GWAS analyses because it benefits both from the precision of association and the robustness of linkage (Meuwissen and Goddard, 2007). Interestingly, LDLA and LDLAb showed similar LRT profiles (except on OAR5 for IgGst). Both methods are based on the same model: a within-family QTL effect that provides the most important features of the differences between MBB and RMN breeds, and paternal and maternal haplotypic effects. The additional information provided by association, therefore, comes from the RMN haplotypes. Indeed, only 4 F_1 sires were available, thus providing 4 different MBB haplotypes segregating at the most, and this did not contribute much to the likelihood. In turn, this leads to few differences in the test statistics of the 2 models, with more degrees of freedom for LDLAb. Subsequently, less QTL are declared significant with LDLAb than with LDLA.

Consistent Findings in Three Regions

Both LA- and LD-based analysis resulted in mapping many QTL involved in resistance to *H. contortus*. However, only a few regions exhibited consistent findings, suggesting the actual presence of a QTL: several correlated traits affected by a limited region (OAR5, 12, and 21) and large significance (OAR12 and 21) or simply a large significance on OAR13.

It is interesting to note that a limited region of OAR12 (between 46 and 56 Mbp) was associated with each of the FEC traits, whatever the infection rank by *H. contortus*, and was confirmed in the BC2 population. The QTL affecting FEC were also mapped on OAR12 in a Merino flock (Beh et al., 2002) and in the free-living Soay sheep population (Beraldi et al., 2007). In the latter, the QTL was mapped between CSSM3 and MCMA52 located at 40.4 Mbp and 60.8 Mbp, respectively (Archibald et al., 2010), which is in good agreement with the results reported herein. Taken together these findings suggest that OAR12 is a key player affecting the excretion of *H. contortus* eggs. However, LD in our population did not help refine the location of this QTL. The GWASb confirmed that MBB haplotypes had the most favorable effect, whereas haplotypes linked to susceptibility segregated in the RMN breed (data not shown). It is therefore difficult to determine whether only one gene underlies this QTL region or whether several genes are present.

By contrast, both OAR5 and OAR13 benefited from LD information; the LRT statistic was affected by the addition of RMN haplotypes to the model, hence greatly enhancing the original signals observed by LA analysis. On OAR5, several peaks were present between 85 and 93 Mbp and were consistently associated with FEC, IgGst, or PCV. Here again, it is difficult to determine whether a unique gene with pleiotropic effects is present. By contrast, the 2-Mbp region between 70.2 and 72.2 Mbp on OAR13 was significantly associated with FEC34t in every analysis. This is a particularly interesting finding because the region is small. Given the QTL effect on FEC34t and the relatively large correlations with other traits (e.g., WBt and PCV2), one would have expected these latter traits to be affected by this region.

First Report of a Significant QTL Affecting Pepsinogen Concentration

The BC2 lambs were measured for pepsinogenemia that is associated with abomasal mucosa disruption during hemonchosis (Simpson, 2000). Interestingly, at reinfection, smaller FEC values were observed in lambs that had greater pepsinogenemia as described by Terefe et al. (2007), who compared the immune responses to *H. contortus* infection in MBB and RMN lambs. Corroborating these findings, a QTL region was mapped on OAR21 for both Peps34t and FECt_a. In addition, multiple associations with different LD-based analyses highlighted a 1-Mbp region that was significantly associated with pepsinogen concentration upon both first and second infections. As far as we know, this is the first significant QTL associated with pepsinogen concentration during GIN infection in sheep. Gutierrez-Gil et al. (2009) measured this trait in a Churra ewe granddaugh-

ter design but did not detect any QTL that segregated for this trait. This difference in results may be explained by our experimental design because infection, as well as the time point of pepsinogen measurement, was strictly controlled. In addition, the GWASb model showed that MBB haplotypes were not found in the RMN breed and were associated with greater pepsinogen concentrations. Interestingly, the bovine *pepsinogen A* locus (*PGA5*) has been positioned between 37,514,983 and 37,525,605 bp (Archibald et al., 2010) and hence constitutes the most obvious functional candidate gene. Several other genes related to the immune response are located in the vicinity of this gene: CD5 known to participate in the selection and activation of both B and T cells and CD6, which is involved in the inhibition of macrophage apoptosis. Additionally, Dominik et al. (2010) found a QTL for the variation in blood eosinophil counts following a 1st infection by *T. colubriformis* in the same region: between CSRD72 and BMS1948 located, respectively, at 40.3 and 51.2 Mbp on OAR21 (Archibald et al., 2010). Even if worm genera are different, eosinophils are known to be key players in the rejection of worms (Meeusen et Balic, 2000; Terefe et al., 2009). In addition, Terefe et al. (2007) reported greater eosinophil counts and greater pepsinogen concentrations in MBB purebred animals, whatever the time of infection by *H. contortus*. Together, these independent findings suggest that a relatively small region on OAR21 might affect blood eosinophils, pepsinogen concentration, and FEC during GIN infection.

Conclusion

In summary, this is one of the most extensive descriptions of *H. contortus* infection in sheep and one of the first SNP-based QTL detection studies for *H. contortus* in sheep.

Many QTL of small to moderate effects were found, and 4 main regions on OAR5, 12, 13, and 21 were identified. The LD phase persistence between the 2 breeds constitutive of the backcross population was small, and some haplotypes were breed-specific. Therefore, a specific model was implemented to cluster haplotypes according to their breed origin, while looking for haplotype-trait associations. This model fitted better reality and outperformed the simple association model that did not take breed origin into account. The chromosomes coming from the MBB breed were usually associated with the most favorable effects. The QTL mapped on OAR5 and OAR12 were found with microsatellite and SNP markers, the confidence intervals being narrower in the latter case.

We report the first QTL affecting pepsinogen concentration, with the region of greatest significance mapping precisely to the *PGA5* gene on OAR21. The OAR5 and OAR13 showed QTL with large effects or pleiotropic effects or both that could not be matched to any known functional candidate genes. The OAR12 remains an interesting candidate because a 10-Mbp region was consistently involved upon both first and second infections. Further work would be required to dissect the underlying genetic mechanisms.

LITERATURE CITED

Amarante, A. F. T., P. A. Bricarello, R. A. Rocha, and S. M. Gennari. 2004. Resistance of Santa Ines, Suffolk and Ile de France sheep to naturally acquired gastrointestinal nematode infections. Vet. Parasitol. 120:91–106.

Archibald, A. L., N. E. Cockett, B. P. Dalrymple, T. Faraut, J. W. Kijas, J. F. Maddox, J. C. McEwan, V. Hutton Oddy, H. W. Raadsma, C. Wade, J. Wang, W. Wang, and X. Xun. 2010. The sheep genome reference sequence: A work in progress. Anim. Genet. 41:449–453.

Aumont, G., L. Gruner, and G. Hostache. 2003. Comparison of the resistance to sympatric and allopatric isolates of Haemonchus contortus of Black Belly sheep in Guadeloupe (FWI) and of INRA 401 sheep in France. Vet. Parasitol. 116:139–150.

Beavis, W. D. 1994. The power and deceit of QTL experiments: Lessons from comparative QTL studies. Pages 250–256 in The Proceedings of the Forty-Ninth Annual Corn & Sorghum Industry Research Conference. American Seed Trade Assoc., Washington, DC.

Beh, K. J., D. J. Hulme, M. J. Callaghan, Z. Leish, I. Lenane, R. G. Windon, and J. F. Maddox. 2002. A genome scan for quantitative trait loci affecting resistance to Trichostrongylus colubriformis in sheep. Animal Genetics 33:97–106.

Beraldi, D., A. F. McRae, J. Gratten, J. G. Pilkington, J. Slate, P. M. Visscher, and J. M. Pemberton. 2007. Quantitative trait loci (QTL) mapping of resistance to strongyles and coccidia in the free-living Soay sheep (*Ovis aries*). Int. J. Parasitol. 37:121–129.

Bishop, S. C., and C. A. Morris. 2007. Genetics of disease resistance in sheep and goats. Small Ruminant Res. 70:48–59.

Bovenhuis, H., and R. J. Spelman. 2000. Selective genotyping to detect quantitative trait loci for multiple traits in outbred populations. J. Dairy Sci. 83:173–180.

Cezar, A. S., G. Toscan, G. Camillo, L. A. Sangioni, H. O. Ribas, and F. S. Vogel. 2010. Multiple resistance of gastrointestinal nematodes to nine different drugs in a sheep flock in southern Brazil. Vet. Parasitol. 173:157–160.

Churchill, G. A., and R. W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. Genetics 138:963–971.

Cole, J. B., P. M. VanRaden, J. R. O'Connell, C. P. Van Tassell, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and G. R. Wiggans. 2009. Distribution and location of genetic effects for dairy traits. J Dairy Sci. 92:2931–2946.

Coop, R. L., W. D. Smith, K. W. Angus, R. B. Graham, S. E. Wright, and F. Jackson. 1985. Effect of *Ostertagia ostertagi* on lamb performance and cross resistance to *O. circumcincta*. Res. Vet. Sci. 39:200–206.

Courtney, C. H., C. F. Parker, K. E. McClure, and R. P. Herd. 1985. Resistance of nonlambing exotic and domestic ewes to naturally acquired gastrointestinal nematodes. Int. J. Parasitol. 15:239–243.

Crawford, A. M., K. A. Paterson, K. G. Dodds, C. Diez Tascon, P. A. Williamson, M. Roberts Thomson, S. A. Bisset, A. E. Beattie, G. J. Greer, R. S. Green, R. Wheeler, R. J. Shaw, K. Knowler, J. C. McEwan. 2006. Discovery of quantitative trait loci for resistance to parasitic nematode infection in sheep: I. Analysis of outcross pedigrees. BMC Genomics 7:178.

Darvasi, A., A. Weinreb, V. Minke, J. I. Weller, and M. Soller. 1993. Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. Genetics 134:943–951.

Davies, G., M. J. Stear, M. Benothman, O. Abuagob, A. Kerr, S. Mitchell, and S. C. Bishop. 2006. Quantitative trait loci associated with

parasitic infection in Scottish blackface sheep. Heredity 96:252–258.

Dominik, S., P. W. Hunt, J. McNally, A. Murrell, A. Hall, and I. W. Purvis. 2010. Detection of quantitative trait loci for internal parasite resistance in sheep. I. Linkage analysis in a Romney × Merino sheep backcross population. Parasitology 137:1275–1282.

Dorny, P., and J. Vercruysse. 1998. Evaluation of a micro method for the routine determination of serum pepsinogen in cattle. Res. Vet. Sci. 65:259–262.

Elsen, J. M., B. Mangin, B. Goffinet, D. Boichard, and P. Le Roy. 1999. Alternative models for QTL detection in livestock. I. General introduction. Genet. Sel. Evol. 31:213.

Good, B., J. P. Hanrahan, B. A. Crowley, and G. Mulcahy. 2006. Texel sheep are more resistant to natural nematode challenge than Suffolk sheep based on faecal egg count and nematode burden. Vet. Parasitol. 136:317–327.

Green, P., K. Falls, and S. Crooks. 1990. Documentation for CRI-MAP. v. 2.4. Washington Univ. School of Medicine, St. Louis, MO.

Gutierrez-Gil, B., J. Perez, L. Alvarez, M. Martinez-Valladares, L. F. de la Fuente, Y. Bayon, A. Meana, F. San Primitivo, F. A. Rojo-Vazquez, and J. J. Arranz. 2009. Quantitative trait loci for resistance to trichostrongylid infection in Spanish Churra sheep. Genet. Sel. Evol. 41:46–**[AU: last page?]**.

Hayes, B., and M. E. Goddard. 2001. The distribution of the effects of genes affecting quantitative traits in livestock. Genet. Sel. Evol. 33:209–229.

Hill, W. G., and A. Robertson. 1968. Linkage disequilibrium in finite populations. Theor. Appl. Genet. 38:226–231.

Hoglund, J., K. Gustafsson, B. L. Ljungstrom, A. Engstrom, A. Donnan and P. Skuce. 2009. Anthelmintic resistance in Swedish sheep flocks based on a comparison of the results from the faecal egg count reduction test and resistant allele frequencies of the beta-tubulin gene. Vet Parasitol. 161:60–68.

Howell, S. B., J. M. Burke, J. E. Miller, T. H. Terrill, E. Valencia, M. J. Williams, L. H. Williamson, A. M. Zajac, and R. M. Kaplan. 2008. Prevalence of anthelmintic resistance on sheep and goat farms in the southeastern United States. J. Am. Vet. Med. Assoc. 233:1913–1919.

Kemper, K. E., D. L. Emery, S. C. Bishop, H. Oddy, B. J. Hayes, S. Dominik, J. M. Henshall, and M. E. Goddard. 2011. The distribution of SNP marker effects for faecal worm egg count in sheep, and the feasibility of using these markers to predict genetic merit for resistance to worm infections. Genet Res. 93:203–219.

Knott, S. A., L. Marklund, C. S. Haley, K. Andersson, W. Davies, H. Ellegren, M. Fredholm, I. Hansson, B. Hoyheim, K. Lundström, M. Moller, and L. Andersson. 1998. Multiple marker mapping of quantitative trait loci in a cross between outbred wild boar and large white pigs. Genetics 149:1069–1080.

Lacroux, C., T. H. Nguyen, O. Andreoletti, F. Prevot, C. Grisez, J. P. Bergeaud, L. Gruner, J. C. Brunel, D. Francois, P. Dorchies, P. Jacquiet. 2006. *Haemonchus contortus* (Nematoda: Trichostrongylidae) infection in lambs elicits an unequivocal Th2 immune response. Vet. Res. 37:607–622.

Lander, E. S., and D. Botstein. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199.

Lander, E., and L. Kruglyak. 1995. Genetic dissection of complex traits: Guidelines for interpreting and reporting linkage results. Nat Genet. 11:241–247.

Legarra, A., and R. L. Fernando. 2009. Linear models for joint association and linkage QTL mapping. Genet. Sel. Evol. 41:43.

Lynch, M., and B. Walsh. 1998. Genetics Analysis of Quantitative Traits. Sinauer Associates, Sunderland, MA.

Mandonnet, N., M. Bachand, M. Mahieu, R. Arquet, F. Baudron, L. Abinne-Molza, H. Varo, and G. Aumont. 2005. Impact on productivity of peri-parturient rise in fecal egg counts in Creole goats in the humid tropics. Vet. Parasitol. 134:249–259.

Mandonnet, N., V. Ducrocq, R. Arquet, and G. Aumont. 2003. Mortality of Creole kids during infection with gastrointestinal strongyles: A survival analysis. J. Anim. Sci. 81:2401–2408.

Marshall, K., J. F. Maddox, S. H. Lee, Y. Zhang, L. Kahn, H.-U. Graser, C. Gondro, S. W. Walkden-Brown, and J. H. J. van der Werf. 2009. Genetic mapping of quantitative trait loci for resistance to *Haemonchus contortus* in sheep. Anim. Genet. 40:262–272.

Matika, O., R. Pong-Wong, J. A. Wooliams, and S. C. Bishop. 2011. Confirmation of two quantitative trait loci regions for nematode resistance in commercial British terminal sire breeds. Animal 5:1149–1156.

McRae, A. F., J. C. McEwan, K. G. Dodds, T. Wilson, A. M. Crawford, and J. Slate. 2002. Linkage disequilibrium in domestic sheep. Genetics 160:1113–1122.

Meadows, J. R., E. K. Chan, and J. W. Kijas. 2008. Linkage disequilibrium compared between five populations of domestic sheep. BMC Genet. 9:61.

Meeusen, E. N., and A. Balic. 2000. Do eosinophils have a role in the killing of helminth parasites? Parasitol. Today 16:95–101.

Melchinger, A. E., H. F. Utz, and C. C. Schön. 1998. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. Genetics 149:383–403.

Meuwissen, T. H. E., and M. E. Goddard. 2007. Multipoint identity-by-descent prediction using dense markers to map quantitative trait loci and estimate effective population size. Genetics 176:2551–2560.

Ooijen, J. 1992. Accuracy of mapping quantitative trait loci in autogamous species. Theor. Appl. Genet. 84:803–811.

Pérez-Enciso, M., and L. Varona. 2000. Quantitative trait loci mapping in F2 crosses between outbred lines. Genetics 155:391–405.

Piepho, H. P. 2001. A quick method for computing approximate thresholds for quantitative trait loci detection. Genetics 157:425–432.

Raynaud, J. P. 1970. Etude de l□efficacité d□une technique de coproscopie quantitative pour le diagnostic de routine et le contrôle des infestations parasitaires des bovins, ovins, équins et porcins. Ann. Parasitol. 45:321–342.

de Roos, A. P., B. J. Hayes, R. J. Spelman, and M. E. Goddard. 2008. Linkage disequilibrium and persistence of phase in Holstein-Friesian, Jersey and Angus cattle. Genetics 179:1503–1512.

Sargison, N. D., F. Jackson, D. J. Bartley, D. J. Wilson, L. J. Stenhouse, and C. D. Penny. 2007. Observations on the emergence of multiple anthelmintic resistance in sheep flocks in the south-east of Scotland. Vet. Parasitol. 145:65–76.

Simpson, H. V. 2000. Pathophysiology of abomasal parasitism: is the host or parasite responsible? Vet J. 160:177–191.

Stear, M. J., M. Parkand, S. C. Bishop. 1996. The key components of resistance to *Ostertagia circumcincta* in lambs. Parasitol. Today 12:438–441.

Terefe, G., C. Lacroux, O. Andreoletti, C. Grisez, F. Prevot, J. P. Bergeaud, J. Penicaud, V. Rouillon, L. Gruner, J. C. Brunel, A. Legarra, J. M. Elsen, J. Bouix, R. Rupp, and C. R. Moreno. 2007. Immune response to *Haemonchus contortus* infection in susceptible (INRA 401) and resistant (Barbados Black Belly) breeds of lambs. Parasite Immunol. 29:415–424.

Terefe, G., C. Lacroux, F. Prévot, C. Grisez, J. P. Bergeaud, C. Bleuart, P. Dorchies, G. Foucras, and P. Jacquiet. 2009. Eosinophils in Haemonchus contortus-infected resistant and susceptible breeds of sheep: Abomasal tissue recruitment and in vitro functional state. Vet. Parasitol. 165:161–164.

Traversa, D., B. Paoletti, D. Otranto, and J. Miller. 2007. First report of multiple drug resistance in trichostrongyles affecting sheep under field conditions in Italy. Parasitol Res. 101:1713–1716.

Yazwinski, T. A., L. Goode, D. J. Moncol, G. W. Morgan, and A. C. Linnerud. 1980. *Haemonchus contortus* resistance in straightbred and crossbred Barbados Blackbelly sheep. J Anim. Sci. 51:279–284.

APPENDIX 1: EXPRESSION OF THE TRUE CORRELATION OF VARIABLES MEASURED ON A SELECTIVE SET OF DATA

Assuming Variable A is measured on the complete data set. Because Variables B and C are only measured on a selected Group G from A, the aim of this appendix is to estimate the true correlations between A, B, and C using observations in the Group G: A_G, B_G, and C_G.

In general, the correlation between Variables A and B is:

$$r(A,B) = \beta(A,B)\frac{\sigma_A}{\sigma_B}$$

where σ_A and σ_B are the SD of Variables A and B and $\beta(A,B)$ is the regression coefficient of A on B.

Because the regression coefficient is not affected by the data selection:

$$r(A,B) = r(A_G,B_G)\frac{\sigma_{B_G}.\sigma_A}{\sigma_{A_G}\sigma_B}$$

The unknown parameter in this equation is σ_B. $\acute{o}^2_B = \acute{o}^2_{eB} + \hat{a}^2(A,B)\acute{o}^2_A$, σ_{eB} is the residual variance of the regression of B on A. Because this variance is not affected by the selection of A: $\acute{o}^2_{B_G} = \acute{o}^2_{eB} + \hat{a}^2(A_G,B_G)\acute{o}^2_{A_G}$, then $\acute{o}^2_B = \acute{o}^2_{B_G} + \hat{a}^2(A_G,B_G)(\acute{o}^2_A - \acute{o}^2_{A_G})$.

Finally, the correlation between A and B is:

$$r(A,B) = r(A_S,B_S).\frac{\sigma_{B_G}.\sigma_A}{\sigma_{A_G}\sqrt{\left(\sigma^2_{B_G} + \beta^2(A_G,B_G).(\sigma^2_A - \sigma^2_{A_G})\right)}}$$

The determination of the correlation between Variables B and C is complicated by the fact that both variables are measured on the selected Group G.

$$r(B,C) = \frac{\mathrm{cov}(B,C)}{\sigma_B\sigma_C},$$

Assuming D = B + C, then $2\,\mathrm{cov}(B,C) = \acute{o}^2_D - (\acute{o}^2_B + \acute{o}^2_C)$ and $\beta(A,D) = \beta(A,B) + \beta(A,C)$.

$$2\,\mathrm{cov}(B,C) = \sigma^2_{D_G} - \left(\beta^2(A_G,D_G) - \left(\sigma^2_{B_G} - \beta^2(A_G,B_G) + \sigma^2_{C_G} - \beta^2(A_G,C_G)\right)\right)\times(\sigma^2_A - \sigma^2_{A_G})$$

$$\mathrm{cov}(B,C) = \mathrm{cov}(B_G,C_G) - \beta(A_G,B_G)\beta(A_G,C_G)(\sigma^2_A - \sigma^2_{A_G})$$

$$\mathrm{cov}(B,C) = r(B_G,C_G)\sigma_{B_G}\sigma_{C_G} - \beta(A,B_G)\beta(A,C_G)(\sigma^2_A - \sigma^2_{A_G})$$

$$r(B,C) = \frac{r(B_G,C_G)\sigma_{B_G}\sigma_{C_G} - \beta(A,B_G)\beta(A,C_G)(\sigma^2_A - \sigma^2_{A_G})}{\sigma_B\sigma_C}$$

$$r(B,C) = \frac{r(B_G,C_G)\sigma_{B_G}\sigma_{C_G} - \beta(A,B_G)\beta(A,C_G)(\sigma^2_A - \sigma^2_{A_G})}{\sqrt{\left[\sigma^2_{B_G} - \beta(A,B)(\sigma^2_A - \sigma^2_{A_G})\right]\left[\sigma^2_{C_G} - \beta(A,C)(\sigma^2_A - \sigma^2_{A_G})\right]}}$$

APPENDIX 2: TRUE QTL EFFECT UNDER SELECTIVE GENOTYPING

As previously, A is measured on the complete data set, and marker genotyping and B phenotyping were performed only on the Group G: the proportion *p* of the low and high tails of the distribution of A. In this context, the true estimations of the QTL effect in a sire QTL design for A and B were proposed by Bovenhuis and Spelman (2000).

QTL Effect on A under Selective Genotyping on A

Assuming a QTL is segregating in the population, the true additive QTL effect is noted *a*, and the estimated QTL effect in the selective genotyping group G is a_G. Consequently, the estimated QTL effect by Qtlmap software can be corrected following the proposition of by Bovenhuis and Spelman (2000):

$$a = \frac{a_G}{\left(1 + Z_{1-p/2}i_{p/2}\right)}$$

where $z_{1-p/2}$ is the deviation of the truncation point from the mean corresponding to p/2; and $i_{p/2}$ is the selection intensity corresponding to p/2.

QTL Effect on B under Selective Genotyping on A

Assuming the true additive QTL effect on B is noted *b*, and the estimated QTL effect in the selective genotyping group G is b_G. Because the regression coefficient is not affected by the data selection, Bovenhuis and Spelman (2000) showed that:

$$b = b_G - \beta(A,B)\left(\frac{Z_{1-p/2}i_{p/2}a_G}{1 + Z_{1-p/2}i_{p/2}}\right)$$

$\beta(A,B)$ is the regression coefficient of A on B.

6.1.2 Discussion

A full discussion of the results has been provided in the paper. Still, additional questions related to this study remain that could not be included in the paper.

The first question would adress the chosen QTL mapping methodology. The QTLMAP software has been chosen as it is especially dedicated for QTL mapping study in outbred populations and as additional developments required by the use of SNP data (huge amount of data, computation efficiency, association and LDLA analyses) were already or simultaneously developped and implemented by the QTLMAP development team. But other methodologies have been proposed to make the best use of SNP data and LD while looking for QTL. Applying these methodologies to our dataset may confirm the obtained results and could also bring a different perspective on our results. Hence, three additional methods have been applied to the OAR12 data in the frame of another study described in the next section (see section 6.2).

Still focusing on the employed methodology, the BC population only provides a sample from the two pure breeds under study. This is especially true for the MBB breed, as only five grand-sires contributed to the BC flock (this point has been fully discussed in the paper [454]). Therefore, trying to exploit ancestral LD after breed formation could be hampered by this narrow view of the populations. Provided that both breeds underwent different histories, especially in terms of GIN-mediated selection pressure, some regions of their respective genomes may have been differentially selected. To look for potential selective sweep in the MBB breed, a pool of MBB sheep have been genotyped [454] and a selection sweep detection has been performed in collaboration with S. Boitard (Laboratoire de Génétique Cellulaire, INRA Toulouse). Methods and results are fully described in this chapter (see 6.3).

Parallel to the methods-related questions is the biological reality of what has been found in the QTL mapping study. Indeed the genetic structure has only been partly unravelled without that any precise structure for these QTL could be proposed. Interestingly, two overlapping QTL signals affecting FEC have been detected on OAR12 at both first and second challenges. This speaks either for two distinct QTL as maximal LRT positions were 10 Mbp apart, or for a key player chromosomic segment carrying a gene cluster acting each time the host faces *H. contortus*. To gain further knowledge about the OAR12 QTL region a validation study has been undertaken. In addition, the presence of an obvious candidate on OAR21 QTL for pepsinogen concentration variation during GIN infection was concomitantly investigated in the second part of the PhD project. Associated experiments and results for these two research tracks are provided in two

dedicated chapter 7 and 8.

6.2 Testing other methodologies on OAR12 for FEC at first infection

6.2.1 Introduction

In the QTL detection study, different models were applied that varied according to the considered historical LD, *i.e.* recent LD events in the within family analysis or ancestral LD that occurred before (GWAS and LDLA without considering the haplotype breed origin) or after breeds formation (GWASb and LDLAb) [454]. Many other methods exist in human and animal genetics literature.

In the frame of the ANR funded "Rules & Tools" project, performances of various QTL mapping methods have been compared using five real datasets from four livestock species (dairy and beef cattle, sheep, pig and horse). The ovine dataset consisted in the OAR12 genomic data of the back-cross population considered in the QTL mapping study while the considered phenotype was mean FEC at first infection.

Three methods have been compared:

- a regular association analysis methods, corrected by relatedness in the population (EMMAX,[257])
- a method considering haplotypes, clustered by approximate identical by descent probabilities (LDLA, [366])
- a Bayesian method which fits effects of all SNP markers simultaneously (Bayes C π,[205])

Empirical power and accuracy, as well as other properties of the methods, *e.g.* redundancy have been considered.

A preliminary version of the manuscript in preparation is provided to detail the whole research work as well as the obtained results for each of the five datasets under study. Some additional data that could be valuable to the discussion of these results are still missing and discussion is still undergoing. However for the purpose of this PhD manuscript, results for the ovine dataset are discussed.

6.2.2 Paper No. 3 (in preparation)

A real data based comparison of methods for whole-genome QTL fine mapping in four livestock species.

A. Legarra, S. Teyssèdre, G. Sallé, P. Croiseau, S. Allais, M.P. Sanchez, Ricard, A., J.M. Elsen.

INTRODUCTION

In the human and, animal breeding literature, a myriad of methods exist to detect and localise Quantitative Trait Locus (QTL). Methods work on the basis of identity, either by descent or by state (one is a proxy for the other; see for instance Cockerham, 1969; Powell et al., 2010), at either singular Single Nucleotide Locus (SNP) or a short series of them, i.e., a haplotype.

Current state of the art, especially in human genetics, consist in a consecutive series of single-marker tests (e.g. association analysis) more often corrected by stratification or coancestry (Kang et al., 2010; Price et al., 2006; Zhang et al., 2003). However, for certain cases, consideration of a series of consecutive loci can be more powerful or more accurate (Browning et Thompson, 2012) for instance if causal QTL are at low allelic frequency. In addition, recent proposals suggest fitting all markers simultaneously (Meuwissen et al., 2001). A comparison of methods using simulation was done by Sahana et al. (Sahana et al., 2010). There exists little (if any, to our knowledge) published comparisons using real data. It is the purpose of this paper to compare, for five different traits and four species, methods for QTL detection and localisation.

Three methods will be compared: i) a regular association analysis method, corrected by relatedness in the population; ii) a method considering haplotypes, clustered by approximate identical by descent probabilities; and iii) a Bayesian method which fits effects of all SNP markers simultaneously. Empirical power and accuracy, as well as other properties of the methods, e.g. redundancy, will be checked. Details of the methods and data sets are provided below.

MATERIALS AND METHODS

Data description

Four species, including five populations, were studied. One chromosome by specie was studied. These are described next.

Dairy cattle (Bos Taurus). A population of 1221 bulls belonging to the Montbéliarde breed was studied. The population was complex, with several overlapped generations. The trait studied was 305-day milk yield; in fact, each bull was assigned a pseudo-performance, the Daughter Yield Deviation (VanRaden and Wiggans 1991), which amounts to the average performance of their daughters corrected by other effects such as herd and the genetic value of each mother. This bulls were genotyped with the Illumina Bovine SNP50 BeadChip. Markers were discarded based on low call rate, lack of positioning in the genome, or very high Mendelian inconsistency rate. No minimum minor allele frequency (MAF) was imposed. Overall, 43582 SNP markers were used. Chromosome 1 was studied (2854 markers).

Beef cattle (Bos Taurus). This population included 936 Blonde d'Aquitaine bulls genotyped with the Illumina Bovine SNP50 BeadChip, resulting in 43582 SNP markers. The phenotype was meat tenderness (Allais et al., 2010, JAS) and chromosome 7 was studied (1889 markers).

Meat sheep (Ovis Aries). This population was issued from a backcross BlackBelly x Romane, so that three generations were included: F1, Backcross, and backcross x backcross (1067 animals all confounded). The trait studied was a measure of resistance to nematode infestation (FEC12t), that is the mean fecal egg count after 25 and 35 days after challenge

with 10,000 H.contortus larvae at 3 months of age as described in Sallé et al.(Sallé et al., 2012). A number of 42,469 SNPs (Illumina, Inc OvineSNP50 Beadchip) passed the quality controls including, 1424 of them in the studied chromosome 12 (Sallé et al., 2012).

Trotter horse (Equus caballus). A total of 627 French trotter horses, structured in small, "little related" (but related nonetheless) families were scored for incidence of osteochondrosis in hock (Teyssèdre et al., 2012). Data set included 41249 SNP, 2267 of them in chromosome 3, genotyped with Illumina BeadChip EquineSNP50.

Pig (Sus scrofa). Some 656 animals were analysed for the trait "length of carcass" (Varona et al., 2002). A total of 46,865 SNPs were retained, including 1672 in chromosome 17, genotyped with Illumina PorcineSNP60 DNA Analysis Kit.

A brief description of the data is in Table 1.

Methods

Gene mapping analyses in pure (not crossed) livestock populations present the particularity, in contrast to human populations, that *all* individuals in the data set are related, and most often these relationships are known.

We used three different methods, namely, a haplotype-based linkage/linkage disequilibrium method (so-called LDLA), an association analysis method, and a Bayesian method. A quick description is as follows:

LDLA. This method was originally conceived in 2002 by Meuwissen et al. (Meuwissen et al., 2002) and our implementation is as in (Druet et al., 2008). Genotypes are phased first according to family and linkage disequilibrium information (Druet and Georges, 2010). Haplotypes are defined as four consecutive (polymorphic) SNP. These are traced along a population; if no recombination occurs, they are considered as identical and carriers of the same allele at the QTL. As for the founder haplotypes, whose identity by descent cannot be ascertained, they are clustered according to their resemblance. This resemblance is based in alikeness in state according to a simple coalescent model. Finally, a variance component model is fit and maximized by Restricted Maximum Likelihood, either with haplotyic effects (alternative hypothesis) or without them (null bypothesis). The model includes a polygenic effect structured according to a pedigree-based relationship matrix: $\mathbf{y} = \mathbf{1}\mu + \mathbf{Zu} + \mathbf{Th} + \mathbf{e}$,where **y** are phenotypes, **u** and **h** are polygenic and haplotype effects, and **e** are residuals, and where $Var(\mathbf{h}) = \mathbf{H}\sigma_h^2, Var(\mathbf{u}) = \mathbf{A}\sigma_u^2$ are covariances matrices.

The statistic, a likelihood ratio test, is then formed, whose p-value is computed according to (Visscher, 2006), assuming multivariate normality. The process is repeated for each locus in the chromosome of interest, where new incidence matrices **T** and **H** are created.

EMMAX. This method (Kang et al., 2010) is a simple extension of regular regression-based association analysis. In our implementation, instead of using a pedigree-based relationship matrix, we used a whole-genome SNP-based relationship matrix (VanRaden, 2008). At each locus in the chromosome of interest we fit a model $\mathbf{y} = \mathbf{1}\mu + \mathbf{Zu} + \mathbf{S}s + \mathbf{e}$, where u are polygenic effects ($Var(\mathbf{u}) = \mathbf{G}\sigma_u^2$) and *s* is a covariate coded as 0,1,2 for each genotype. Effect *s* was estimated by BLUP using blupf90 (Misztal et al., 2002). A t-statistic was constructed as $t = \hat{s} / s.e.(\hat{s})$.

107

108 *BayesC.* This method (Habier et al., 2011) forms part of a Bayesian family originally
109 conceived for prediction of genetic merit and phenotypes, later extended to map gene
110 locations (Hoggart et al., 2008; Meuwissen et al., 2001), but which has the potential to
111 simultaneously analyze all genome. It includes a set of variable indicators, $\mathbf{d} = \{d_1, \ldots d_n\}$,
112 which indicate if a locus "is" (d_i=1) or "is not" (d_i=0) in the model $\mathbf{y} = \mathbf{1}\mu \sum \mathbf{S}_i d_i s_i + \mathbf{e}$, and
113 provides marginal *a posteriori* inference on **d**, in the form of Bayes Factors (Kass et Raftery,
114 1995; Wakefield, 2009). A key parameter of the model is the amount of SNPs entering into
115 the model, which was fixed at 1/1000. Experimentation with 1/100 or 1/10000 did not change
116 qualitatively the results (data not shown). Gibbs sampling chains were run for 100,000
117 iterations using GS3 (Duchemin et al., 2012). All SNPs were included in the model although
118 only the chsomosome of interest was studied later. Two reported statistics were used. The first
119 was the posterior probability of a locus being "in" the model, $\hat{d}_i$ (which goes from 0 to 1). The
120 second was the Bayes Factor, which corresponds to the increase of evidence from the prior to
121 the posterior, that is : $BF_i = \hat{d}_i / (1/1000)$. The Bayes Factor is akin (but not identical to) the
122 likelihood ratio (Kass et Raftery, 1995).
123

124 We used chromosome-wise Bonferroni correction to infer rejection thresholds for p-
125 values, and a uniform threshold of 0.5 for the posterior probability $\hat{d}_i$ (this value implies that
126 it is more likely than a locus "is" in the model than the opposite).
127

128 In order to get a rough idea on the complexity (or redundancy) of each combination
129 method-data set, we used two ideas. The first is to use a number of equivalent tests, using
130 Geyer's "effective sample size", where the genome can be seen as a Markov Chain (Geyer et

Thompson, 1995). This allows having a rough idea of how many independent tests would provide the same amount of information. The second estimate is Goddard et al. (2012) “number of independent chromosomal segments”, which can be inferred from marker information.

RESULTS

Manhattan plots for each population are presented in Figures 1 to 5. Summarizing the results is complicated and for clarity purpose we have chosen to present the followings.

Empirical power. We present the number of tests above the nominal threshold described above (Table 2). Because we don't know the actual number of QTLs, it has to be regarded with caution. It can be seen that LDLA has the higher nominal power, whereas BayesC seems short of power. EMMAX is somewhere in the middle.

Agreement. The objective of QTL mapping is not only to detect, but to provide locations which molecular geneticists can further explore. Hence, consistency of QTL mapping was checked (Table 3). . Large agreement across methods can be seen from this table and corresponding figures (Figures 1 to 5), especially for cattle populations. For instance, dairy cattle data agrees on a window of 2 cM around position 65 cM while beef cattle related QTL map two positions around 98 and 32 cM Major inconsistencies were found in the pig population, maximal test statistics being obtained in four different regions around 7, 18, 30 and 40 cM (Table 3).

Measure of redundancy. These are presented in Table 4. In general, LDLA tends to be rather redundant, whereas EMMAX and BayesC are less so. This is because haplotypes are in strong correlation from one position to the next, whereas SNP effects are not. In addition, BayesC estimators are marginalised over the rest of loci. However, Goddard's number of equivalent segments gives a different picture, with similar figure for all species except for sheep, which are a back-crossed population.

REFERENCES

Browning, S.R. et E.A. Thompson. 2012. Detecting rare variant associations by identity-by-descent mapping in case-control studies. Genetics. 190:p.1521-1531. Available: Consulté juillet 10, 2012.

Druet, T., S. Fritz, M. Boussaha, S. Ben-Jemaa, F. Guillaume, D. Derbala, D. Zelenika, D. Lechner, C. Charon, D. Boichard, et al. 2008. Fine mapping of quantitative trait loci affecting female fertility in dairy cattle on BTA03 using a dense single-nucleotide polymorphism map. Genetics. 178:p.2227-2235. Available: Consulté juillet 18, 2012.

Druet, T. et M. Georges. 2010. A hidden markov model combining linkage and linkage disequilibrium information for haplotype reconstruction and quantitative trait locus fine mapping. Genetics. 184:p.789-798. Available: Consulté juillet 18, 2012.

Duchemin, S.I., C. Colombani, A. Legarra, G. Baloche, H. Larroque, J.-M. Astruc, F. Barillet, C. Robert-Granié et E. Manfredi. 2012. Genomic selection in the French Lacaune dairy sheep breed. Journal of dairy science. 95:p.2723-2733. Available: Consulté mai 5, 2012.

Geyer, C.J. et E.A. Thompson. 1995. Annealing Markov Chain Monte Carlo with applications to ancestral inference. Journal of the American Statistical Association. 90:p.909-920.

Habier, D., R.L. Fernando, K. Kizilkaya et D.J. Garrick. 2011. Extension of the bayesian alphabet for genomic selection. BMC bioinformatics. 12:p.186. Available: Consulté mai 5, 2012.

Hoggart, C.J., J.C. Whittaker, M. De Iorio et D.J. Balding. 2008. Simultaneous analysis of all SNPs in genome-wide and re-sequencing association studies. PLoS genetics. 4:p.e1000130. Available: Consulté juillet 10, 2012.

Kang, H.M., J.H. Sul, S.K. Service, N.A. Zaitlen, S.-Y. Kong, N.B. Freimer, C. Sabatti et E. Eskin. 2010. Variance component model to account for sample structure in genome-wide association studies. Nature genetics. 42:p.348-354.

Kass, R.E. et A.E. Raftery. 1995. Bayes Factors. Journal of the American Statistical Association. 90:p.773-795.

Meuwissen, T.H., B.J. Hayes et M E Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics. 157:p.1819-1829. Meuwissen, T.H.E., A. Karlsen, S. Lien, I. Olsaker et Mike E Goddard. 2002. Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping. Genetics. 161:p.373-379.

Price, A.L., N.J. Patterson, R.M. Plenge, M.E. Weinblatt, N.A. Shadick et D. Reich. 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nature genetics. 38:p.904-909. Sahana, G., B. Guldbrandtsen, L. Janss et M.S. Lund. 2010. Comparison of association mapping methods in a complex pedigreed population. Genetic epidemiology. 34:p.455-462.

Sallé, G., P. Jacquiet, L. Gruner, J. Cortet, C. Sauvé, F. Prévot, C. Grisez, J.P. Bergeaud, L. Schibler, A. Tircazes, et al. 2012. A genome scan for QTL affecting resistance to Haemonchus contortus in sheep. Journal of animal science.

Teyssèdre, S., M.C. Dupuis, G. Guérin, L. Schibler, J.M. Denoix, J.M. Elsen et A. Ricard. 2012. Genome-wide association studies for osteochondrosis in French Trotter horses. Journal of Animal Science. 90:p.45-53.

VanRaden, P.M. 2008. Efficient methods to compute genomic predictions. Journal of dairy science. 91:p.4414-4423.

Varona, L., C. Ovilo, A. Clop, J.L. Noguera, M. Pérez-Enciso, A. Coll, J.M. Folch, C. Barragán, M.A. Toro, D. Babot, et al. 2002. QTL mapping for growth and carcass traits in an Iberian by Landrace pig intercross: additive, dominant and epistatic effects. Genetical research. 80:p.145-154. A

Visscher, P.M. 2006. A note on the asymptotic distribution of likelihood ratio tests to test variance components. Twin research and human genetics: the official journal of the International Society for Twin Studies. 9:p.490-495.

Wakefield, J. 2009. Bayes factors for genome-wide association studies: comparison with P-values. Genetic epidemiology. 33:p.79-86.

Zhang, S., X. Zhu et H. Zhao. 2003. On a semiparametric test to detect associations between quantitative traits and candidate genes using unrelated individuals. Genetic epidemiology. 24:p.44-56.

TABLES

Table 1: Basic statistics of the data.

	Dairy	Beef	Sheep	Horse	Pigs
No. animals	1221	936	1067	627	656
Trait	Milk yield	Meat tenderness	Fecal egg count	Ostechondrosis score in the hock	Length of carcass
Chromosome investigated	1	7	12	3	17
No. SNPs available	2854	1889	1424	2267	1672
Length in cM	161	112	79	119	60
Recent N_e[a,b]	77	114	50	115	61
Ancient N_e[a,c]	1270	1279	2373	493	8182

[a]: Ne, Effective population size; [b], inferred from neighbour markers; [c], inferred from markers 40-60 cM apart.

Table 2: Number of significant “hits” obtained for the three analyses in each available population.

	Dairy	Beef	Sheep	Horse	Pigs
LDLA*	20	0	29	6	0
EMMAX*	0	2	3	1	0
BayesC**	1	1	0	0	0

* Above chromosome-wise Bonferroni threshold

** Above 0.5 in posterior probability

Table 3. Positions of the two major peaks obtained for the three methodologies in each of the five populations under study

			LDLA	EMMAX	BayesC
Dairy cattle	*1st peak*	**Position**[a]	66.09	70.09	64.09
		Test statistic[b]	*7.44*	*3.81*	*2.91*
	2nd peak	**Position**	64.9	64.93	65.71
		Test statistic	*7.13*	*3.23*	*2.32*
Beef cattle	*1st peak*	**Position**	98.4	98.7	98.7
		Test statistic	*4.37*	*5.89*	*2.91*
	2nd peak	**Position**	98.69	32.77	32.77
		Test statistic	*4.32*	*4.93*	*2.47*
Horse	*1st peak*	**Position**	105.05	105.88	105.88
		Test statistic	*5.66*	*5.29*	*2.68*
	2nd peak	**Position**	105.13	102.03	68.81
		Test statistic	*5.21*	*4.62*	*2.23*
Sheep	*1st peak*	**Position**	51.2	13.83	51.19
		Test statistic	*5.66*	*5.26*	*2.56*
	2nd peak	**Position**	51.31	10.61	10.61
		Test statistic	*5.62*	*4.49*	*2.41*
Pig	*1st peak*	**Position**	40.93	42.32	30.38
		Test statistic	*3.11*	*2.91*	*1.53*
	2nd peak	**Position**	18.13	7.1	7.1
		Test statistic	*3.09*	*2.73*	*1.12*

[a], given in cM

[b], -log10(p-value) for LDLA and EMMAX analyses and –log10(Bayes Factor) for BayesC methodology

239 Table 4: Number of “independent” tests

	Dairy	Beef	Sheep	Horse	Pigs
LDLA	237	183	153	978	380
EMMAX	2323	1082	770	968	1672
BayesC	2390	1889	1108	1303	1509
Goddard’s M,* 1 chromosome	81	97	42	72	91

240 *Number of equivalent chromosomal segments: 1/mean(r^2) over 1 chromosome.

241

FIGURES

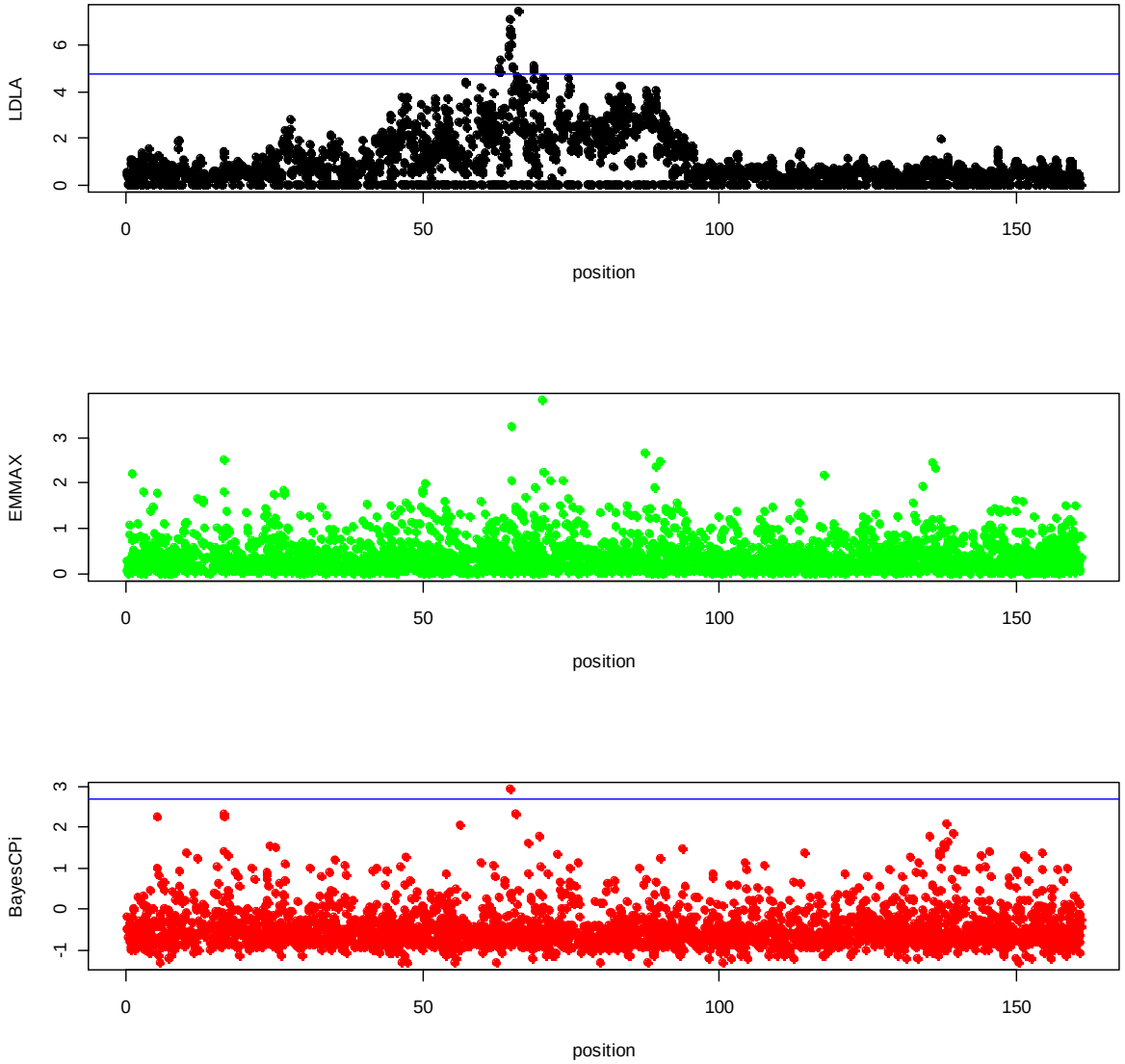
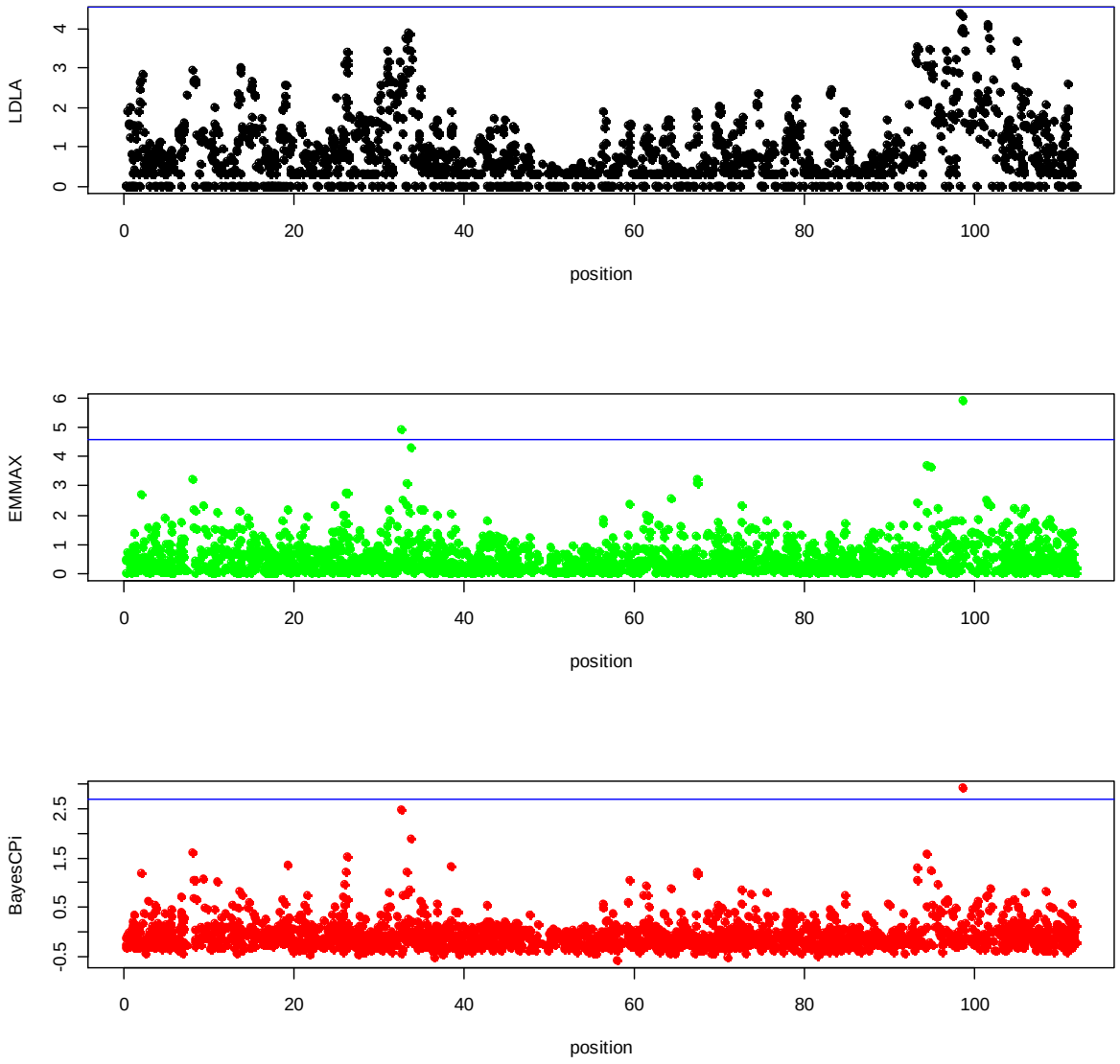


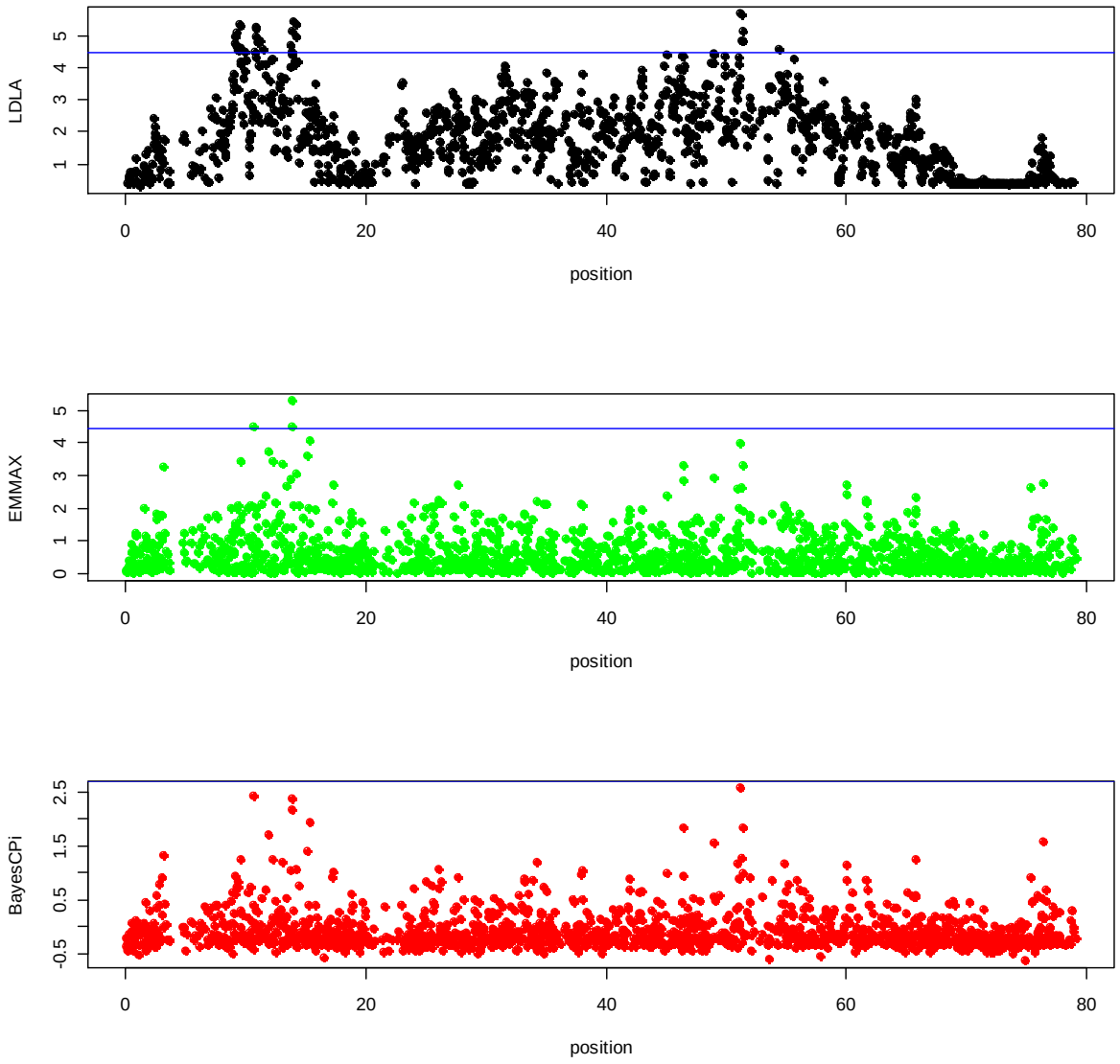
Figure 1. Scan of chromosome 1 in dairy cattle. Y-axis: log10(1/p-value) for LDLA and EMMAX; log10(Bayes Factor) for BayesCPi). X-axis: position along the chromosome (in cM). The blue line (if any) is the rejection threshold.



248

249 Figure 2. Scan of chromosome 7 in beef cattle. Y-axis: log10(1/p-value) for LDLA and
250 EMMAX; log10(Bayes Factor) for BayesCPi). X-axis: position along the chromosome (in
251 cM). The blue line (if any) is the rejection threshold.

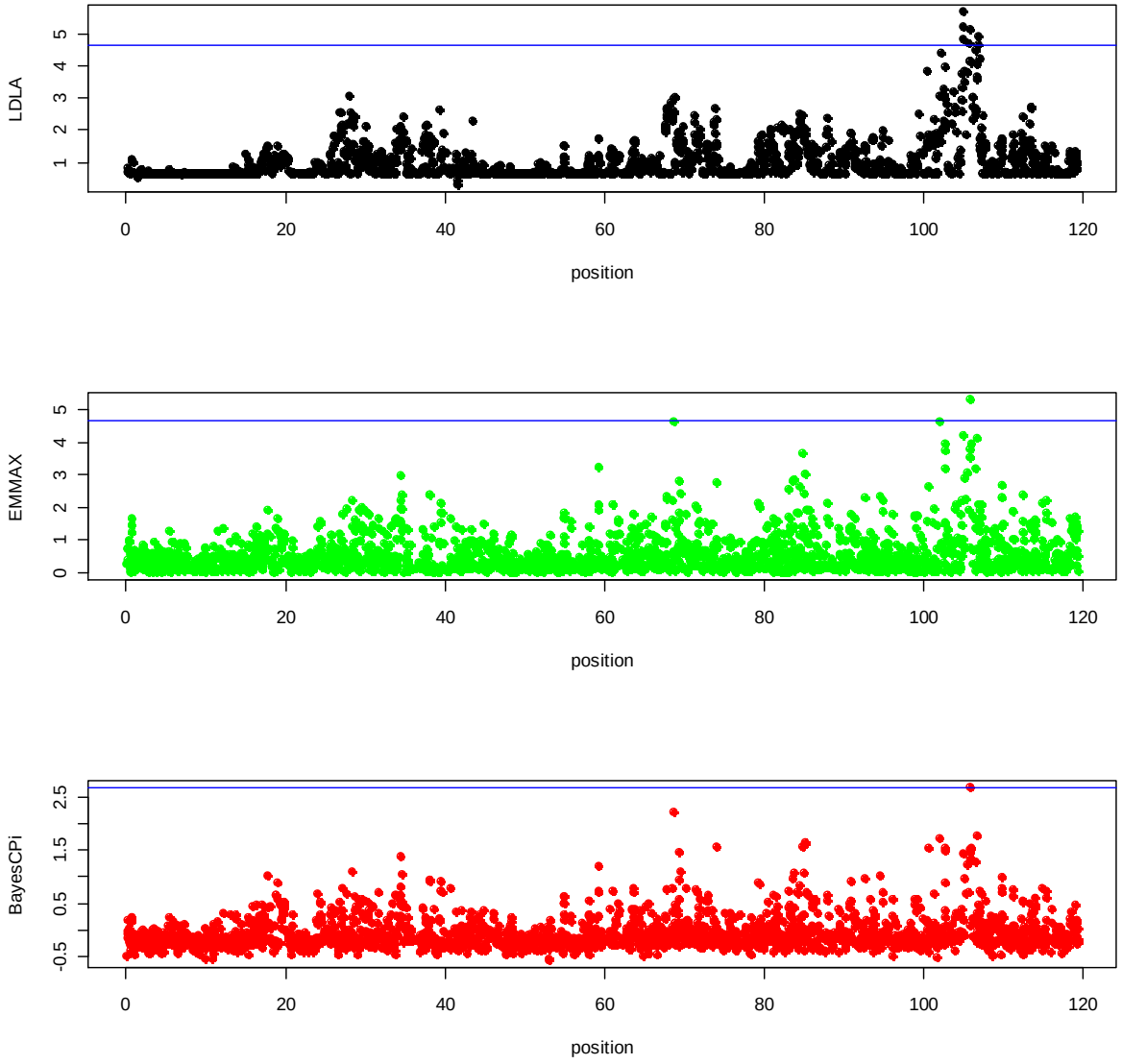
252



253

254 Figure 3. Scan of chromosome 12 in sheep. Y-axis: log10(1/p-value) for LDLA and
255 EMMAX; log10(Bayes Factor) for BayesCPi). X-axis: position along the chromosome (in
256 cM). The blue line (if any) is the rejection threshold.

257



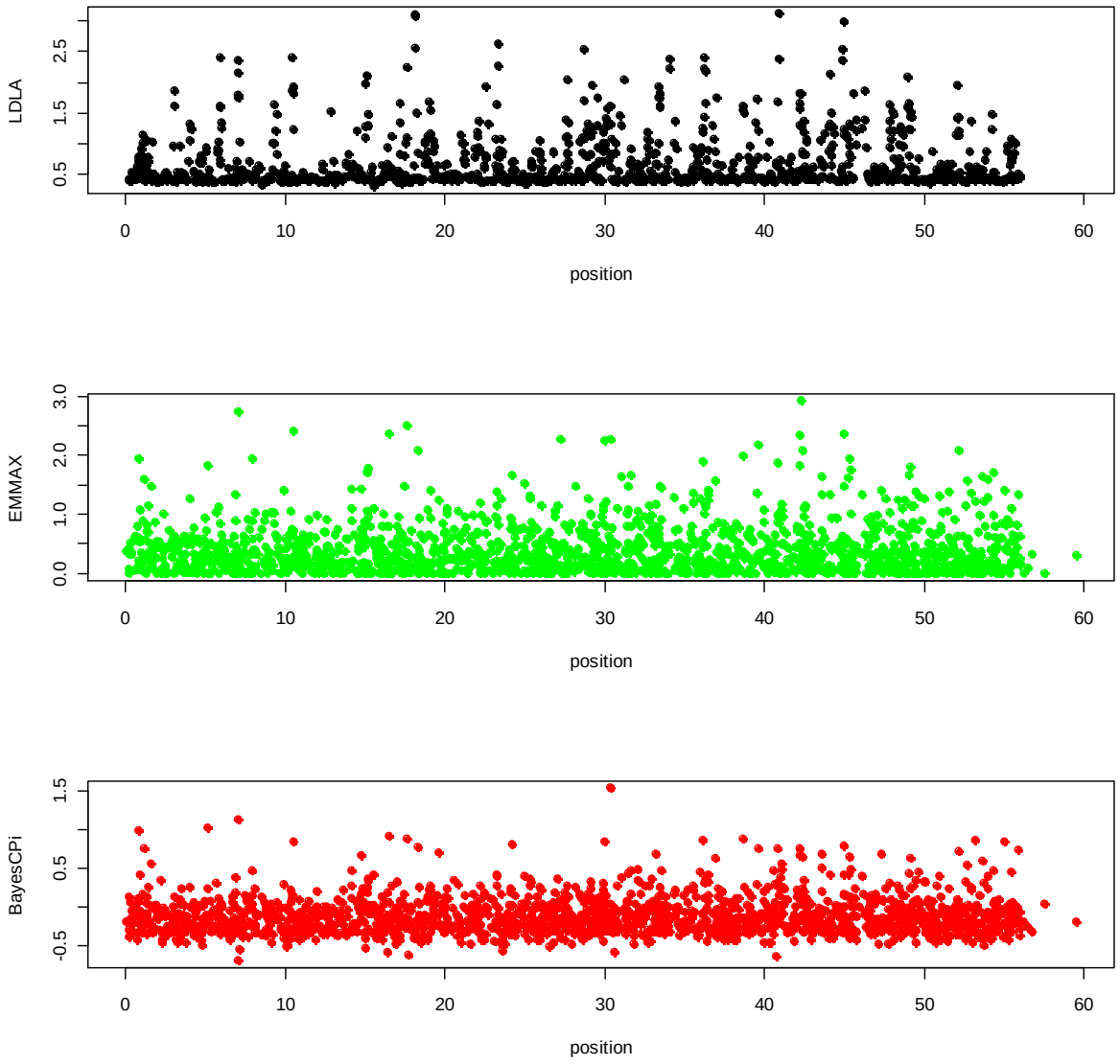
258

259 Figure 4. Scan of chromosome 3 in horses. Y-axis: log10(1/p-value) for LDLA and EMMAX;

260 log10(Bayes Factor) for BayesCPi). X-axis: position along the chromosome (in cM). The blue

261 line (if any) is the rejection threshold.

262



263

264 Figure 5. Scan of chromosome 17 in pig. Y-axis: log10(1/p-value) for LDLA and EMMAX;

265 log10(Bayes Factor) for BayesCPi). X-axis: position along the chromosome (in cM). The blue

266 line (if any) is the rejection threshold.

267

268

269

270

271

6.2.3 Discussion about the ovine dataset

In the end, all methods seem to result in similar conclusions with few spurious peaks, if any, arising due to stratification. This is mainly because relationships have been accounted for.

LDLA was generally the most powerful method, whereas EMMAX was the most robust but results show variations according to the considered dataset.

Focusing on the sheep data only, it appears that methods agreed well on fitting the maximal probability of QTL segregating around 50 Mbp and another peak around 10 Mbp. Interestingly, the Meuwissen & *al.*'s LDLA was the only method to put the two most important peaks close to the position found by the QTLMAP software (56 Mbp).

From a graphical aspect, it is worth that likelihood profiles are very similar between the QTLMAP software analysis (method by Legarra & Fernando, [307]) and the LDLA from Meuwissen & al. [366] while BayesC and EMMAX methodologies both result in a thiner QTL profile. In a more schematic representation, LDLA landscape looks like moutains chain whereas two other methods provides flatter landscape with some chimneys in the middle.

In this way, LDLA associated likelihood profiles look like a superposition of the association study over the within family knowledge basis. In strong relationship to the number of available meiosis, within-family analysis provides a large signal centered around 50 Mbp, that extends over more than 20 Mbp. Association of haplotypes at the population scale adds likelihood points to the corresponding position that result in some "chimneys" on the top of the large "moutain". The two other methods considered do not involve an explicit within-family component. Population structure is accounted for by the genomic similarity summarized in the 50K SNP markers. This information is formalized by a genomic relationship matrix in the EMMAX method that is used to correct the estimated SNP effect. For the bayesian methodology, the correction for stratification occurs through the simultaneous consideration of all SNP which in turn also identifies different genomic patterns. In the end, it seems that EMMAX and BayesC π buffer the background noise of the non QTL-associated SNP. This discussion is purely speculative as it is precisely not known where the QTL is/are. Simulation studies could help better understanding differences between different methods.

The origin of the considered chromosomal segments could influence QTL detection results. In the QTL detection study By Sallé & *al.* [454] the breed clustering of haplotypes lead to a dramatic increase of the likelihood on OAR5 and OAR13. In the study presented in this section, none of the analysis considered this parameter hence resulting in looking for QTL

appeared before breed diverged. In the employed methodologies, stratification seems to be correctly accounted for thanks to the whole genome information that averages relationships between individuals. But, even if BC lambs carry one quarter of a MBB genome and 75% of a RMN genome on average, there is 50% chance that the two alleles come from different breeds at a given position. Predicting how the three investigated methodologies behave with this parameter does not seem to be straighforward and simulations should be a good way to investigate this point further.

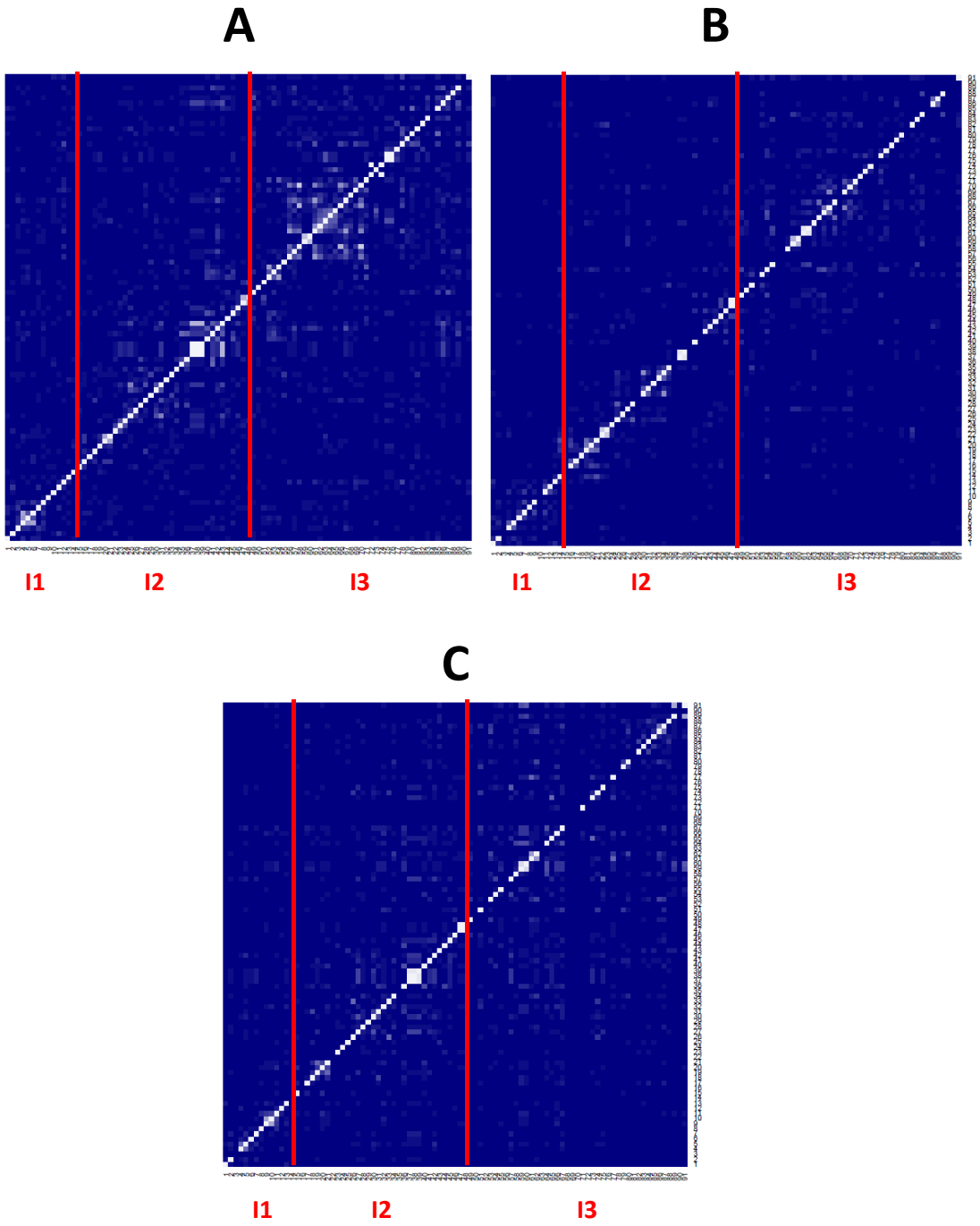
Considering both the results from LA and LD-based analyses, it is worth noticing that the 10 Mbp region that is ranked as the second peak (LDLA, Bayes C π) or even represents the two highest test statistics (EMMAX) did not appear in the within family analysis (figure 2A in [454], see section 6.1.1). Hence, this peak could be a spurious signal that arises due to a strong linkage disequilibrium with the previous identified QTL around 50 Mbp [417].

To mine the first hypothesis, LD between markers included in the chromosome region that exhibited a LRT equal to (LRTmax - 2LOD) was estimated in pure breeds and in BC progenies. Resulting relationship were plotted as a heatmap displayed on figure 6.1. In the end 91 SNP were retained, 13 within 9 and 12 Mbp, 35 within 25 and 41 Mbp and 43 SNP in the 46 to 56.6 Mbp interval. As shown previously [454], moderate to high value of LD were observed between SNP relatively far apart in the MBB breed and several LD hotspots could be observed between SNP located around 49 and 51 Mbp in this breed (figure 6.1 A). These hotspots did not hold in the RMN population neither long range LD (figure 6.1). The pattern of the BC population was intermediate between the two pure breeds populations. Considering the BC pattern, there is little chance that a spurious peak between SNP far apart, typically 10 and 56 Mbp could arise by strong LD.

Another explanation of the 10 Mbp likelihood peak observed on OAR12 could be a second QTL signal arising thanks to the linkage disequilibrium information contained in the RMN chromosome. To investigate this point, two QTL detection were performed either considering (full model) or not (restricted model) the maternal haplotypes. The difference between the LRT of two models should provide information about the maternal haplotypes contribution to the likelihood. Resulting profiles for the two models and the difference between LRT are provided in figure 6.2. This figure clearly shows that keeping only the RMN maternal contribution to the likelihood, results in a peak located at 11 Mbp.

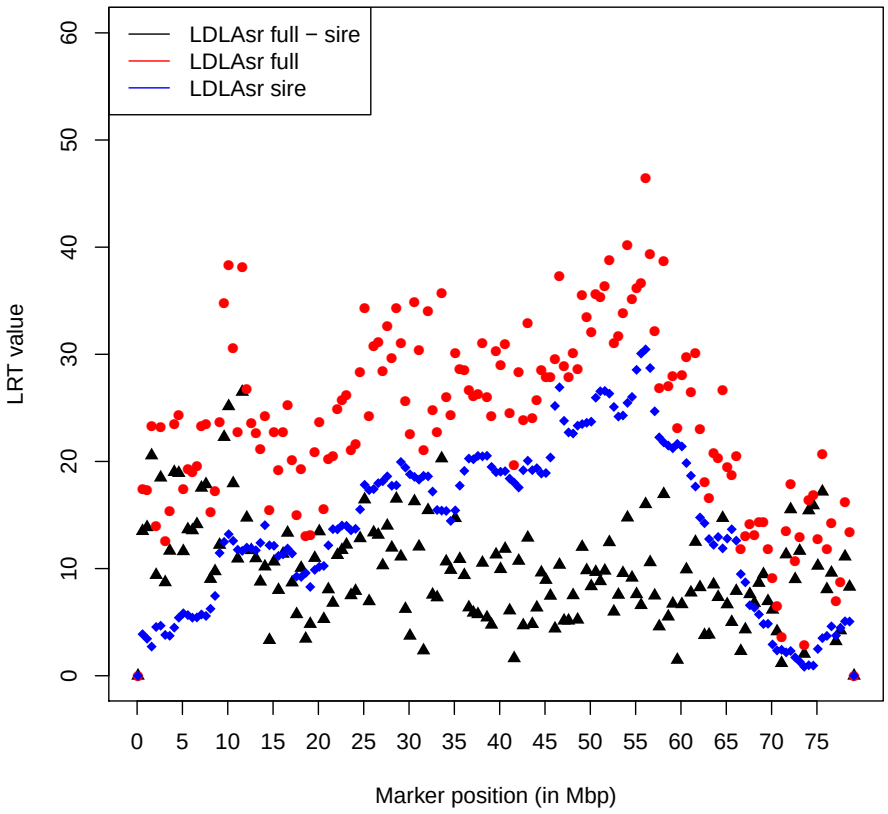
Additional simulation studies are now required for completing this study. Real data will con-

Figure 6.1: LD between SNP with LRT above maximal LRT - 2LOD in the MBB, RMN and BC populations



r^2 coefficient between pairs of every SNPs located within an area exhibiting a LRT value above a (LRTmax-2LOD) value of the GWAS analysis, is plotted as a heatmap. SNPs are clustered according to their location: the I1 group representing SNPs located between 9.5 and 11.6 Mbp, I2 for SNPs located between 25 and 40 Mbp and I3 for SNPs within 46 and 56 Mbp. A, B and C

Figure 6.2: Isolation of a RMN specific QTL at 10 Mbp on OAR12



The likelihood of the full QTL detection model is plotted in red. The reduced likelihood model that does not consider maternal haplotypes is plotted in blue. Black plots are for the difference between the likelihood in the two models

stitute the template before assigning QTL properties to some SNPs (Elsen & Legarra, personal communication). Hence pedigree structure and genotypic data will remain unchanged. Phenotypes will be simulated in relationship to the SNPs chosen as QTL. At the moment, several cases should be considered with varying MAF and LD between SNP, several QTL effect and dominance or not. The number of QTL should vary as well to investigate the effect of epistasis.

6.3 Additional track: looking for selective sweep in pure breeds

6.3.1 Rationale

MBB and RMN breeds have undergone different histories especially, they evolved in different environment with different exposure to GIN. Through natural selection, it may be possible that GIN among other factors have put selection pressure on some regions of the genome in the MBB breed. The reported analysis aims at detecting these regions and to see whether some of them corresponds to any of the detected QTL regions. This analysis has been performed in collaboration with S. Boitard (Laboratoire de Génétique Cellulaire, INRA, Castanet-Tolosan).

6.3.2 Materials and methods

6.3.2.1 Animals and genotypes

Genotypic data have been already described in [454]. Briefly, a sample of 90 MBB sheep structured in nuclear families have been selected to be the most representative of their breed and subsequently genotyped with the 50K ovine SNP chip. The RMN data came from a granddaughter design included in the SheepSNPQTL project.

6.3.2.2 Sweep detection

This work was performed by S. Boitard (INRA, Laboratoire de Génétique Cellulaire). We looked for hard sweep signatures within MBB and RMN using the Hidden Markov Model (HMM) in [59]. In this model, the derived allele frequency at SNP i, denoted Y_i, is taken as the observed state at this SNP. Each SNP i is also assumed to have a hidden state X_i, which can take 3 different values : "Selection", for SNPs that are very close to a swept site, "Neutral", for SNPs that are far away from any swept site, and "Intermediate" for SNPs in between. These three values are associated with different allele frequency distributions.

The "Neutral" allele frequency distribution is estimated using all SNPs in the genome, assuming most of them have indeed evolved under neutrality. Allele frequency distributions in states "Intermediate" and "Selection" are deduced from this "Neutral" distribution using the derivations in [390], and are typically more skewed towards extreme allele frequencies. The hidden states X_i form a Markov chain along the genome with a per bp probability p of switching state, so that close SNPs tend to be in the same hidden state.

Under this HMM, the most likely sequence of hidden states can be predicted from the sequence of observed states using the Viterbi algorithm. Each set of consecutive SNPs with predicted state "Selection" is called a sweep window. Besides, applying the backward-forward algorithm to the same HMM provides, for each SNP i, the posterior probability q_i of hidden state "Selection". This probability quantifies the evidence of a sweep at each SNP. Since it is oten close to 1 in "Selection" windows, it is generally expressed in log scale, by $-log(1-q_i)$.

To avoid using related individuals, which may bias the allele frequency estimation, we applied the method in MBB using only the 51 founders (*i. e.* removing offsprings and other relatives). In RMN, the pedigree only included 8 founders, so we added 50 maternal chromosomes in order to increase the sample size. These 50 haplotypes were chosen to maximize the genetic diversity observed in RMN. To achieve this objective, we computed the Identity by State distance between all pairs of maternal chromosomes in the pedigree, and then kept the 50 haplotypes maximizing these distances using the hclut() function in R.

To determine which of the two alleles at a given SNP is ancestral, we used results obtained by the ISGC during the Sheep HapMap project (J. Kijas, personal communication). During this project, several animals from ovis species were genotyped using the Sheep SNP50 chip, so alleles that were present in all these species could be considered as ancestral alleles for Sheep. This strategy provided the ancestral allele at 31592 SNPs. For other SNPs, we used a folded allele frequency distribution, i.e allele frequencies Y_i and $1-Y_i$ were considered as the same observed state.

The genome wide type I error of the above method, i.e. the probability that it detects at least one sweep window in a population that has evolved under neutrality, depends on parameter p (see [59] for more details). To control this type I error at level 5%, we generated 100 permutations of the allele frequency vector, and adjusted p in order to detect sweep windows in at most one of these permutations. This lead to $p = 6.67572e-07$ for MBB and $p = 5.72205e-07$ for RMN.

6.3.3 Results

The full results of the sweep detection in MBB are plotted in figure 6.3. No sweep window was detected at level 5%, but several regions with hidden state "Intermediate" (suggestive selection signature) were found.

Noteworthy, one region under selection was found on OAR12 at 42 Mbp, close to the maximal LRT values found for FEC at 2nd infection and within the QTL confidence interval. The region under selection was composed of seven SNP as illustrated on figure 6.4. The five first SNP were fixed in the investigated MBB sample population but still segregated in the RMN population. The sixth SNP broke the sweep signal and the seventh SNP was also fixed.

Interestingly, when fitting the sweep genotype as a fixed effect of the linkage analysis, dramatic fall of the likelihood was observed for FEC at first and second infection (figures 6.5).

Every F1 sire carried the swept MBB allele and the two sires associated to families with significant QTL signal carried the same RMN allele for the 5 first SNP.

This five-SNP region encompasses one gene whose function cannot be obviously related to parasite infection (see figure 6.6).

In the RMN breed, two other suggestive selection signature could be found on OAR6 (at 36 and 37.2 Mbp) and on OAR10 (at 10 Mbp) but none of these regions did correspond to one of the already detected QTL.

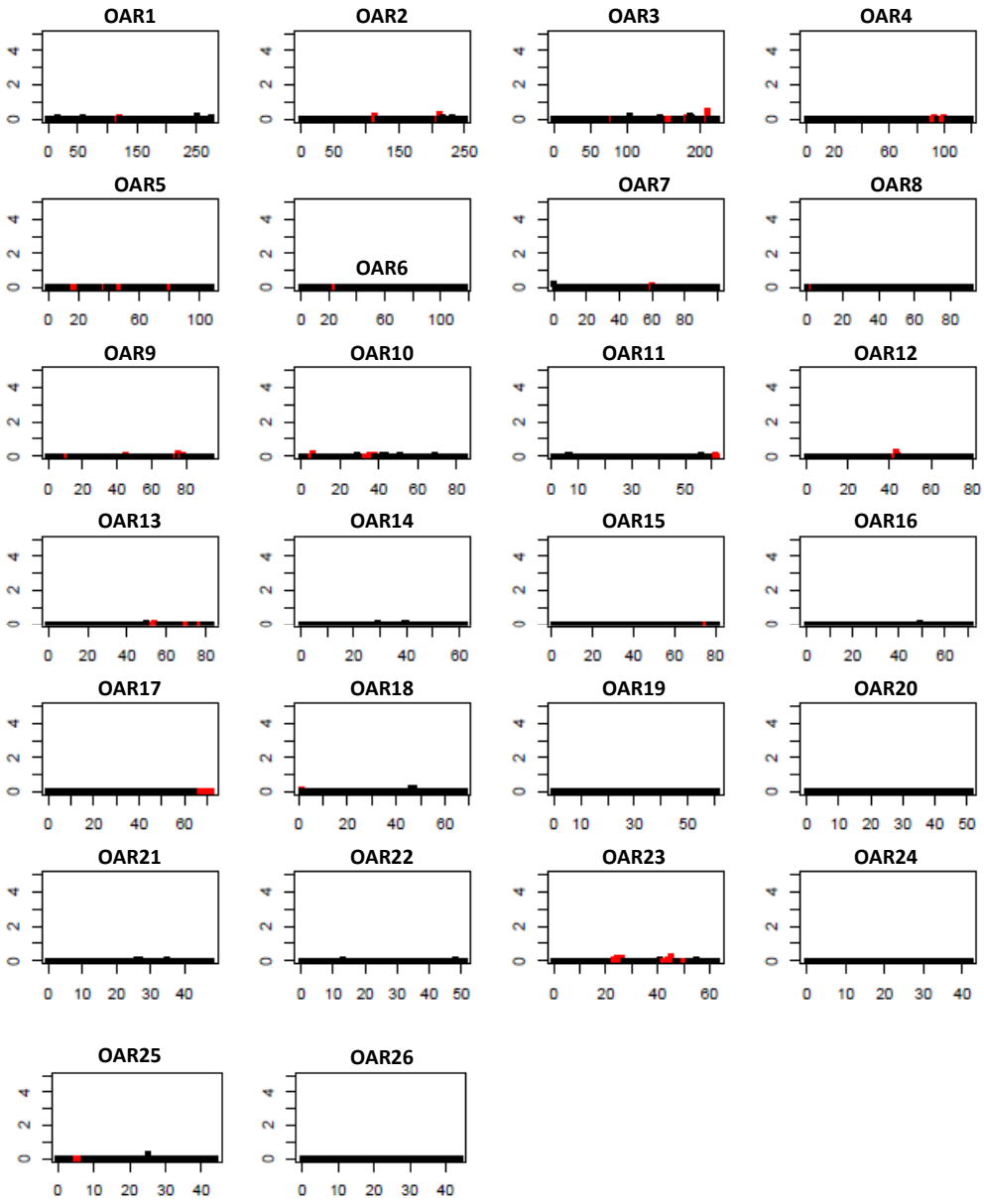
6.3.4 Discussion

MBB and RMN breeds have undergone different evolutive processes. *H. contortus* mediated selection should have been different between the MBB breed that lives in subtropical area where *Haemonchus* is predominant. Such selection can lead to the reduction of allelic frequencies in the vicinity of the genes that are selected. On the contrary, selection can also be diversifying as reported for the MHC locus [426].

In this work, regions of the genome showing reduced homozygosity were looked for in the MBB breed hence following the hypothesis of a directional selection mediated by *H. contortus* infection.

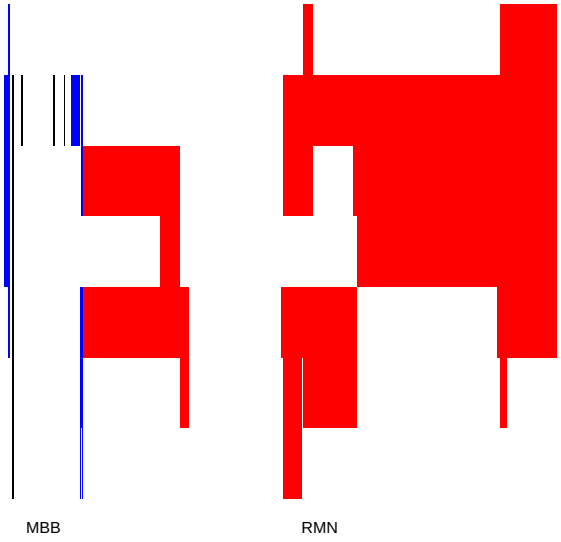
Interestingly, a five-SNP region located within the QTL CI was found to be fixed in the MBB breed whereas these five SNPs were still segregating in the RMN breed. In addition, fitting the sweep genotype as a fixed effect in the QTL analysis greatly reduced the likelihood in the area hence suggesting that this region is in disequilibrium with the QTL.

Figure 6.3: Results of the sweep detection analysis in the MBB breed



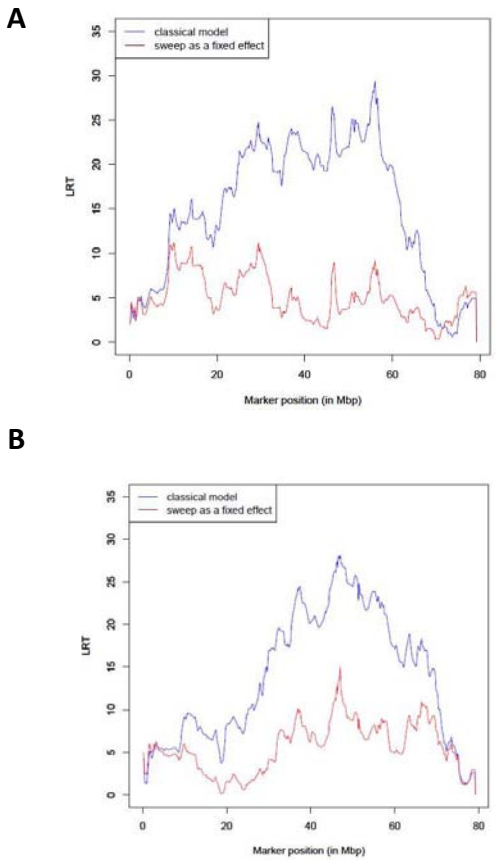
Each box corresponds to a chromosome, where the Y-axis is the probability of sweep occurring and the X-axis represents the SNP positions in Mbp along the chromosome

Figure 6.4: Schematic representation of the sweep region in the MBB and RMN breeds



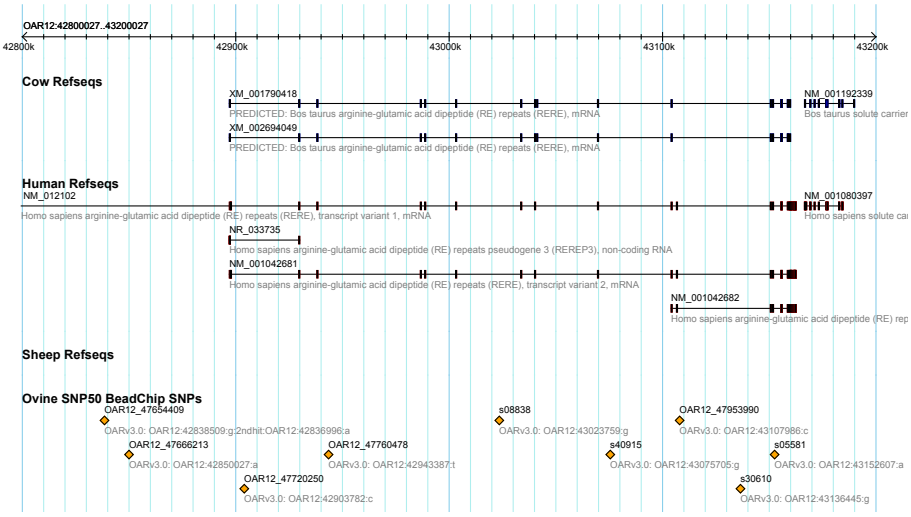
The seven SNP of the sweep are represented by a vertical line (each SNP being a segment of this line, the first SNP being at the bottom), each line corresponding to one individual's genotype. All genotypes are put together along the horizontal axis and grouped by population. Within each line, white segment corresponds to the allele under selection, other alleles being either blue or red for the MBB and RMN populations respectively

Figure 6.5: LA QTL profile after fitting the sweep genotype as a fixed effect



A: Evolution of the likelihood for the linkage analysis performed on FEC at first infection without (blue) or with the sweep as a fixed effect (red); B: Evolution of the likelihood for the linkage analysis performed on FEC at second infection without (blue) or with the sweep as a fixed effect (red);

Figure 6.6: Schematic representation of the sweep region in the MBB and RMN breeds



This snapshot was taken from the CSIRO sheep genome browser version 3 (http://www.livestockgenomics.csiro.au/cgi-bin/gbrowse/oarv3.0/, access the 23^{th} of july 2012

Still, elements are missing to conclude. Firstly, no phenotypes were recorded so that it is difficult to assess whether these regions with particular allelic frequencies are in relationship with GIN infection or not. Indeed, even if GIN may be among the most important selection pressure applied to the MBB breed, breeding has certainly highly contributed to shape the allelic frequency pattern in this breed. For instance, the MBB breed is hairy whereas the RMN breed produces wool. Therefore a difference in allelic frequency between the two breeds could also be translate such morphological differences.

Secondly, the close physical relationship between the maximal likelihood positions of the QTL and the sweep hampers the clear dissection of this region. Indeed a QTL analysis run after correction for the sweep genotype resulted in a dramatic fall of the likelihood. However such phenomenon could be generated by the linkage disequilibrium between the sweep and the "true" QTL. Fitting the sweep genotype as a fixed effect should "absorb" the QTL signal as the clustering based on the sweep genotype or the true QTL genotype may lead to the same BC population partitionning.

Thirdly, the sweep signal position perfectly corresponds to a drop of the LRT (data not shown).

The hypothesis underlying this work is that GIN infection has been responsible for a strong directional selection pressure after breed divergence. Let us say that directional selection can occur in the case of disease resistance (which seems not to be the case with the well-known

MHC example, [426]), the question of how much pressure has been put by GIN on the sheep genome remains. The effective selection pressure exerted by GIN under natural conditions is both unknown and difficult to evaluate. The fewer it is, the longer it takes to select any particular region of the genome and the higher number of recombinations between the QTL and close markers. Therefore the SNP density achieved on the 50K SNP chip may not be sufficient to unravel these mutations as only the closest QTL vicinity will be linked to it. Regarding this question, our populations may not be the best as they have also been suggested to other selection pressure linked to breeding objectives. Evolutionary insights may be gained from the Soay sheep breed that faces parasites in natural conditions [86].

It is also possible that some regions have been common to both breeds. To investigate this hypothesis in a more general way, breeds genotyped by the ISGC consortium have been clustered in groups according to their published status towards GIN infection, *i.e.* rather resistant or not. An analysis was performed that aimed at isolating regions of the genome selected for in the resistant breeds either with a geographical clustering or not (S. Boitard). However no significant signal could be isolated (data not shown). Same kind of results had been found by Kijas & al. (personal communication).

6.4 Conclusion about the QTL detection study

Exploiting data from a 1,000 BC population helped to identify regions of the genome involved in resistance to GIN infection in sheep. A first screening with microsatellites unravelled major regions that were subsequently confirmed with a high-density pangenomic SNP chip. Among these, OAR5, 12 and 13 seem to play a major role in the course of an infection by *H. contortus*.

The OAR12 is particularly interesting as it seems to significantly affect clinical parameters of haemonchosis whatever the rank of infection, hence playing a central role in both naive and immune lambs. The application of several other methodologies on this same region have both confirmed the QTL even if the precise genetic structure of this QTL signal remains unresolved. Intriguingly, a selection sweep has been identified in an independent MBB breed within the QTL confidence interval, 4 Mbp upstream from the maximal LRT value for FEC at reinfection. Noteworthy the same 5-SNPs haplotype was still segregating in another RMN population. No conclusion can be drawn so far, but it can be hypothesized that GIN infection have lead to the selection of this haplotype in the MBB breed.

Adding more SNP markers helped unravelling many other QTL that had not been detected with sparse map among which OAR21 was an original QTL affecting pepsinogen concentration evolution during haemonchosis. The presence of an obvious functional candidate gene makes it a particularly interesting region to be further investigated.

These two main strands of research require additional work that have been conducted during the second part of the PhD project. Results and perspectives are provided in the next sections.

Chapter 7

Functional investigation of the QTL region on OAR12

Résumé

Dans une première analyse QTL, une région du chromosome 12 a été associée au nombre d'oeufs de parasite excrétés par des agneaux naïfs ou immuns durant une infestation expérimentale par *H. contortus*. Par ailleurs, une trace de sélection de cinq SNPs située dans l'interval de confiance du QTL a été identifiée dans une population de Martinik Black-belly. Notre étude a pour objet l'étude fonctionnelle de cette région du génome ovin. Des individus BC ont été choisis sur la base de leur génotype au QTL afin de produire des agneaux croisés BCxBC.

Une première sélection des individus BCxBC a ainsi été réalisée sur la base de leur génotype à la trace de sélection pour étudier l'impact de cette région sur la résistance à *H. contortus*. La comparaison des groupes allèliques obtenus a montré un effet significatif sur la fertilité des vers femelles.

Après que la sélection des animaux BCxBC ait eu lieu, une analyse d'association a été réalisée dans la population parentale BC. Cette analyse a montré qu'une région de quatre SNPs était associée de manière significative à l'intensité d'excrétion d'oeufs durant l'infestation. Tirant parti de ce résultat, les animaux BCxBC ont pu être regroupés à nouveau sur la base de leur génotype à ces 4 SNPs. Cette nouvelle comparaison a non seulement confirmé les résultats obtenus précédemmentla fertilité des vers, mais elle a également mis en évidence une différence significative sur l'intensité d'oeufs excrétés et l'hématocrite des agneaux au cours de l'infestation.

Une analyse d'expression de gènes candidats fonctionnels a montré une plus forte imprégnation

Th-2 de la muqueuse abomasale des animaux porteurs d'un allèle de résistance au QTL par rapport aux porteurs d'un autre allèle associé à la sensibilité à *H.contortus*. Cependant aucun candidat fonctionnel n'a pu être validé.

Summary

In a previous QTL mapping experiment, a region of OAR12 was associated to fecal egg count during *H. contortus* infection in both naive and primed lambs. In addition, a selection sweep of five SNPs located within the confidence interval has been identified in a Martinik Black-belly population. Our study aimed at investigating the functional properties of this QTL. BC sheep were selected based on their QTL genotype to produce BCxBC lambs.

A first selection of BCxBC lambs was based on their genotype at the selection sweep to investigate its putative role on the resistance to *H. contortus*. The comparison of performances of each allelic group demonstrated a significant impact of female worms fertility.

An association study performed in the BC population after the BCxBC creation highlighted a 4-SNP region. Exploiting this result, BCxBC lambs were clustered again into groups based on their allele at the same 4-SNP region. This comparison not only confirmed the findings obtained with the selection sweep, but also resulted in significant differences in FEC output and hematocrit drop under experimental haemonchosis. A gene expression study indicated a higher Th-2 environment in the abomasal mucosa of the individuals carrying one QTL allele associated to resistance in comparison to carriers of another unfavorable allele. However no functional candidate genes has been validated so far.

7.1 Introduction

Anthelmintic resistance is a burning issue throughout the world, especially in small ruminant nematodes [261]. Breeding animals with a better ability to cope with gastro-intestinal nematodes has been proposed as an alternative strategy [49]. It has been estimated that genetics explained 20 to 40% of the observed variation in response to GIN infection [49]. Besides, this genetic variability has been exploited in breeding programs in Australia and New-Zealand [263].

From a functional point of view, resistance to GIN infection mostly relates to the mounting of an efficient Th-2 biased immune response driven by the IL4, IL5 and IL13 cytokines [288, 507].

This immune response is thought to affect regulate female worm fertility [489, 288]. An innate component that regulates the worm burden itself has also been involved, but this component of the response seems to be less heritable [489, 482]. Mast cells are responsible for an immediate hypersensitivity reaction, while eosinophils reduce the establishment of larvae [506]. Besides eosinophils recruitment has been shown to differential between resistant and susceptible breeds [508]. In addition, recent findings suggest that lectins could contribute to entrap worms into mucus sheath, that subsequently facilitate their elimination [165, 166, 316].

Numerous studies have unraveled QTLs affecting fecal egg count (FEC) or other patho-physiological parameters relative to nematode infection [38, 42, 101, 108, 342]. The use of the recently released ovine-specific DNA SNP chip showed that resistance to nematodes follows the infinitesimal model: an infinitesimal number of genes with weak effect and some limited regions explaining more of the genetic variation [270].

Still, it is of interest to map and decipher the genetic architecture of this trait as it will improve both the genetic selection by directly targeting the genes of interest through marker-associated selection, and the knowledge about underlying mechanisms. However, few QTL if any have been dissected to the underlying Quantitative Trait Nucleotide (QTN) [552, 14].

In a previous QTL mapping study for resistance to *Haemonchus contortus*, Sallé *et al.* found five QTL of greater interest on OAR5, 7, 12, 13 and 21 that affected Fecal Egg Count (FEC) and other finer patho-physiological parameters measured in a 1,000 Martinik Black-belly x Romane (MBB x RMN) back-cross (BC) lambs population [454]. The QTL region on OAR12 was particularly interesting as it was constantly associated to FEC in primed and naïve lambs. In addition, a selection sweep lying close to the maximal likelihood position of the QTL has been unraveled in a Martinik Black-belly population.

The purpose of this study was two-fold. The first objective was to confirm the actual presence of a QTL and to dissect further its architecture by comparing progenies selected for carrying particular alleles associated to resistance or susceptibility and to investigate the putative role of the sweep region. These offsprings were obtained by crossing BC individuals together (BCxBC). The second objective was to investigate functional properties of the discovered QTL region, by an exhaustive parasitological and hematological data collection, as well as functional candidate genes comparison of expression.

7.2 Material and methods

7.2.1 Animal material

7.2.1.1 Marker-assisted matings

Animal selection was based on the results of the linkage analysis performed in the back-cross population [454]. This within-family analysis demonstrated that chromosomes carried by two F1 sires exerted opposite effect on FEC at first infection, *i.e.* the MBB allele being associated to a reduction in eggs output [454].

To investigate the biological properties of this QTL region, a marker-assisted mating was performed to produce lambs carrying particular combination of QTL alleles, *i.e.* two alleles associated to resistance/susceptibility, or one of each (see figure 7.1). For doing so, 70 BC animals (21 males and 49 females) were available for selection, as well as 81 BCxBC crosses already produced in 2008 and 2009 (35 males and 46 females). All these animals were genotyped with the ovine 50K SNP chip (Illumina Inc, San Diego, CA).

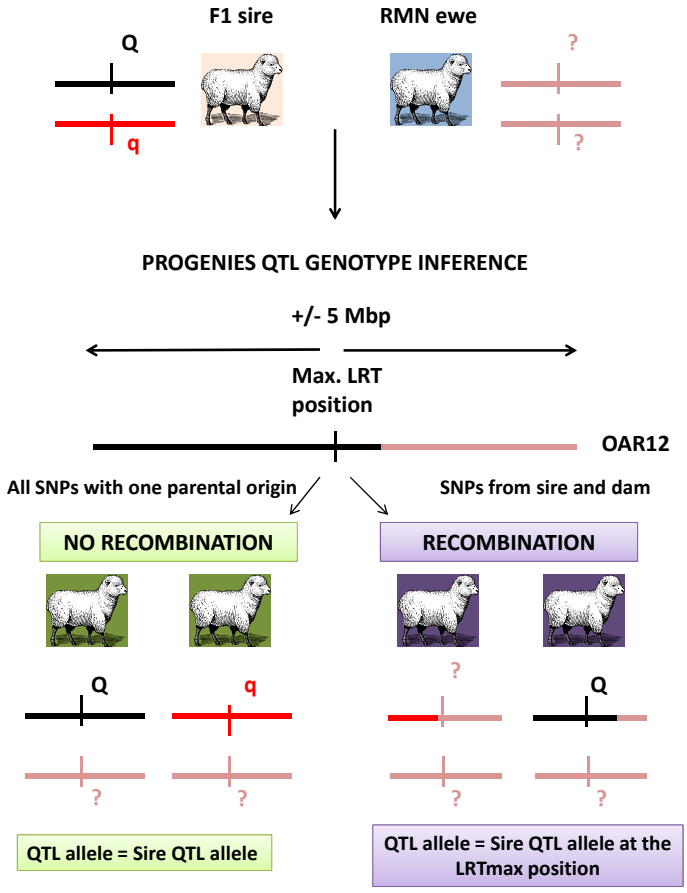
At the time of mating, results from the linkage analysis were the only available knowledge, hence not permitting to pinpoint the QTL. Thus, two 10-Mbp-wide windows centered at the two maximal LRT positions of the QTL affecting FEC at first and second infection (47 and 56 Mbp for mean FEC value at first and second infection respectively) were traced from pure breed grand-parents to BC lambs and BCxBC crosses. This width was chosen as a compromise between increasing the chance of actually targeting the real gene(s) underlying the QTL signal, and avoiding recombination events that make tracing alleles along pedigree more complex.

Chromosomes were reconstructed using the LinkPHASE software [133]. Subsequently, the parental origin of the considered chromosomic segment was inferred as follows (figure 7.1). If less than five informative SNPs (SNP being heterozygous in sire and homozygous in progeny) were found within the first 20 SNPs or the last 20 SNPs of the considered interval, it was checked that no recombination took place in the 100 up- or downstream markers respectively. On average, a chromosome segment was declared inherited without recombination if all informative SNP (42% on average) came from the same parental chromosomes.

In case of a recombination occurring within the 10-Mbp region, the inherited parental allele at the maximal LRT position served as a proxy for the progeny QTL allele.

In the end, QTL genotype was inferred for each available progeny, based on two 10-Mbp regions surrounding the maximal LRT positions of the QTL for FEC at first and second infection.

Figure 7.1: Marker-assisted matings of BC sheep



After chromosome reconstruction with the LinkPHASE software [133], the parental origin of each SNP is determined. If no recombination occurred, QTL allele of progeny is the same as the parental one. In case of recombination, the QTL allele of the parent at the maximal LRT position was considered as the progeny QTL allele.

According to the inherited QTL allele from their sire, each progeny was fitted a QTL index being the estimated QTL allelic substitution effect. The index calculated for the two considered QTL at 47 and 56 Mbp were summed to produce a total QTL12 index. According to this index, animals with the lowest (favorable) and the highest (unfavorable) values were chosen for mating. Female flock was also selected following this scheme.

Finally, two sires carrying QTL resistance alleles were mated to 20 and 15 females of the same genotypic group. Only one "susceptible" male out of the two selected for mating was successful at mating the 47 selected susceptible females. After weaning, 130 out of the 138 born lambs were available for the experiment at the La Sapinière INRA experimental unit.

7.2.1.2 Selection of BCxBC progenies based on the sweep region

At 1 month of age, every born lamb was genotyped with the Illumina 50K SNP chip (Illumina Inc., San Diego, CA) and genotypes were edited as described in [454]. Their chromosomes were reconstructed as already described (last subsubsection 7.2.1.1). Following this work-flow, too few animals carried simultaneously two intact resistant (n=10) or two susceptible 10-Mbp QTL alleles (n=7) of the F1 ancestors. Hence, BCxBC were sorted following the hypothesis that the sweep signal detected in the MBB breed was a putative contributor to the QTL likelihood (see 6.3).

The selection sweep allele fixed in the MBB breed was considered as associated to more resistant phenotypes. On the contrary, the RMN sweep allele carried by the two sires contributing the most to the QTL was thought to confer susceptibility. Three groups of BCxBC lambs were constituted carrying either two MBB sweep allele (denoted SW_{MBB}), two RMN alleles (denoted SW_{RMN}) or one MBB allele (SW_U).

After selection, 61 lambs remained for experiment. These animals were transported from the La Sapinière experimental farm to the Langlade experimental unit for ease of experimental sampling.

7.2.1.3 Second selection of BCxBC progenies based on the association mapping analysis

Subsequent to the within family analysis and the first selection of BCxBC lambs, an association analysis was performed in the BC population [454]. This analysis consisted in testing the effect of a 4-SNP haplotype window on the trait of interest, at every 0.01 Mbp. Haplotypes

whose frequency was below 1% were discarded to limit standard error of the estimation. In addition, breed origin of the haplotype was taken into account, so that two identical-by-state (IBS) haplotypes were considered different if their breed origin was different. In the end, a 4-SNP haplotype positioned at 56.1 Mbp was declared significantly associated to FEC at first infection [454]. Each haplotype effect was estimated.

Exploiting this result, BCxBC lambs could be sorted again into three groups according to the 4-SNP haplotype they carried. A "$4SNP_R$" group was defined as animals carrying one of the alleles with most favorable effect and did not carry alleles associated to unfavorable effect. Conversely, the $4SNP_S$ individuals were chosen as carrying one of the allele associated to susceptibility without any allele associated to low-FEC. Remaining animals were considered as "unknown" ($4SNP_U$).

7.2.2 Experimental infection

7.2.2.1 Infection procedure

Lambs were kept indoor during the whole experiment thus remaining totally worm-free before their infection. Upon arrival, the 61 lambs selected based on their QTL allele, were given a Vecoxan ND treatment (diclazuril, 1 mg/kg bodyweight, Janssen) at the recommended dose to prevent any coccidiosis outbreak. They were subsequently left indoor for a one-month acclimation period, before receiving another anthelmintics drenching treatment to remove any *Strongyloides spp.* (Oramec, 0.25 mg/kg bodyweight, Merial). After checking that no strongyle eggs were excreted, 44 lambs were infected orally with 10,000 infective L3 larvae of the *H.contortus* "Humeau" strain [288]. The remaining 17 control lambs were not challenged. For practical purposes, control lambs were sacrificied two days after the challenge took place, whereas the 44 other infected animals were euthanized at 30 and 31 dpi. Euthanasia was performed by a veterinary surgeon with a lethal intra-venous injection of embutramide (T61, 6 mL/50 kg bodyweight, Intervet). Local INRA procedures for the care of experiment animals were applied throughout the experiment.

7.2.2.2 Pathophysiological measurements and tissue sampling

Intra-rectal collection of feces was performed every three days from 18 days post infection (dpi) until 30 dpi for FEC counting following the McMaster method modified by Raynaud [428]. These traits were denoted **FEC18**, **FEC21**, **FEC24**, **FEC27** and **FEC30**.

Blood samples were collected just before infection, at 14, and 27 dpi. Samples were processed by the Sysmex XT-2000iV hematology analyzer (calibrated for sheep), hence providing a complete screening of hematological parameters (indiced 0, 14 and 27 for samples taken before, at 14 and 27 dpi respectively). In addition to red cells and reticulocytes count (denoted **ret**), white blood cells counts were obtained, *i.e.* lymphocytes (**lymph**), monocytes (**mono**), neutrophils (**neut**), eosinophils (**eo**), basophils (**baso**). Hematocrit was also determined (**hct**).

After sacrifice, draining abomasal lymph nodes were sampled and weighted as well as a patch of the abomasal mucosa. These samples were subsequently stored at -20°C in RNAlater (Ambion, USA).

Abomasal contents and washings were collected and put into absolute alcohol. Worm burden (**WB**) was determined using 10% of the total abomasal content. The lengths of 35 intact adult female worms were determined and averaged (denoted **FL**) for each lamb. To determine the number of eggs *in utero*, 20 additional female worms were digested into a bleaching mixture (40 mL of Milton agent diluted into 160 mL of distilled water) and 10% of the resulting mixture was sampled for eggs counting using an optical microscope (denoted **FF**).

7.2.3 Gene expression measure

7.2.3.1 Total RNA extraction and cDNA amplification

Total mRNA from abomasal fundic mucosa and draining lymph nodes of the sampled animals was extracted following the commercial RNeasy Mini Kit (Qiagen). The RNA quality of the recovered RNA was monitored by A260/A280 spectrophotometry. RNA were subsequently reverse transcripted to cDNA with a Reverse Transcriptase commercial kit (Invitrogen).

7.2.3.2 qPCR analyses

In a previous study [316], some genes had been found differentially expressed between the MBB and the RMN breed, either in abomasal mucosa (GAL15, ITLN2, TFF3) or in draining lymph nodes (OX40, CXCL14, CCL16) or both (IL4, IL13, $TNF\alpha$, $IFN\gamma$). The relative expression of each of these genes was therefore tested in BCxBC lambs. Gene expression before and after infection was compared between two allelic groups.

In addition, the annotated genes the closiest to the 4-SNP region were retrieved from the third assembly of the ovine genome (livestockgenomics.csiro.au/cgi-bin/gbrowse/oarv3.0/). Biological process associated to these genes were retrieved on the NCBI gene website (ncbi.nlm.nih.gov/gene/).

For expression analysis, primers were designed for these particular genes using the primer 3 NCBI website and the bovine transcript as a template. Secondary structures were looked for on the Mfold website (http://mfold.rna.albany.edu/) [581] and selected primer sequences were blasted against the third version of the ovine genome to ensure specificity of their target.

The qPCR was performed with three replicates per sample. Gene expression of five reference genes specific of each tissue was measured. Their gene-wise stability value was estimated as reported in [527] and most stable genes were kept for subsequent analysis. Differential expression was tested following the DDCt method [319]. The cycle time (Ct) value of the gene of interest was corrected by the average level of reference genes expression (dCt). dCt values of the infected animals were corrected by the average dCt values of the corresponding susceptibility control group as follows:

$$ddCt = (Ct_{infected_resistant} - average(Ct_{control_resistant})) - (Ct_{infected_susceptible} - average(Ct_{control_susceptible}))$$

Normality was checked using the Shapiro-Wilk test implemented in the R software (R software, http://CRAN.R-project.org/), and any outlier (+/- 3 standard deviations from mean) was eliminated (n=1).

Fold change in gene expression between considered groups was computed as 2^{-ddCt} [319].

Subsequently, a Wilcoxon test was applied to determine any significant difference between the compared groups, *i.e.* $p < 0.05$. The complete data processing was performed using an homemade R script (R software, http://CRAN.R-project.org/).

7.2.4 Statistical analyses

7.2.4.1 Transformations applied to phenotypes

Normality of the phenotype distribution was checked with the UNIVARIATE procedure of the SAS software. Strong departures from normality were found for FEC that were corrected by a fourth root transformation.

7.2.4.2 Genomic merit estimation

Resistance to nematodes is known to be polygenic and some other major QTL did segregate in this population [454]. Therefore, a genomic value evaluating the effect of the genomic background of each individual was estimated with the Bayes C genomic selection method [205] im-

plemented in the GS3 software [134]. Briefly, this bayesian method analyzes the whole genome simultaneously and determines each SNP contribution to the trait of interest.

Trait considered for analysis was the mean FEC at first infection, corrected for environmental effects of the whole BC and BCxBC population (n=1200 individuals) with 10,000 iterations and burning of the first 200 iterations. The genomic data for gEBV estimation consisted in the 25 autosomes other than OAR12. Indeed, including this latter chromosome was expected to partially sweep the genetic variance explained by the QTL into the gEBV.

7.2.4.3 Testing for statistical differences between allelic groups

The effect of the genotypic group was estimated by the PROC MIXED procedure implemented in SAS (SAS 2001, Cary NC). Usually encountered environmental effects were considered as fixed effects and the computed gEBV was fitted to the model as a covariate to account for the effect of the rest of the genome.

For hematological parameters, basal value of the considered parameters (indexed by 0) were considered as covariate to account for potential inter-individual variation before the beginning of the experiment.

7.3 Results

7.3.1 Overview

A wide range of phenotypes have been measured on the BCxBC lambs. A tabular description of the two clustering performed on the 44 BCxBC lambs is provided in table 7.1. The 44 infected lambs were roughly equally distributed in the different QTL allelic groups, while their estimated gEBV were equivalent (p=0.94 and 0.5 for the sweep-based and 4-SNP based clustering, table 7.1).

Table 7.1: Summary of the allelic groups

	Sweep region			**4-SNP region**		
Group	SW_{MBB}	SW_{RMN}	SW_U	$4SNP_R$	$4SNP_S$	$4SNP_U$
No. Lambs	11	15	18	16	16	12
avg. gEBV[a]	-0.14	-0.13	-0.15	-0.15	-0.15	-0.12

a:genomic breeding value

Basic statistics of parasitological traits are listed in table 7.2 and main features of hematological parameters are provided in table 7.3.

Measured FEC were in a high-range of values with a mean FEC30 of 17,700 eggs/g (table 7.2). Average FEC showed high variance. This was also observed in WB, some animals carrying less than 500 worms while other showed more than 50% establishment (table 7.2).

Table 7.2: Basic statistics of parasitological traits measured in BCxBC lambs

Variable	**Mean**	**Std**	**Minimum**	**Maximum**
FEC18 (eggs/g)	358	795	0	4,150
FEC21 (eggs/g)	2,205	2,978	0	12,500
FEC24 (eggs/g)	5,069	4,250	100	20,600
FEC27 (eggs/g)	7,093	5,302	200	25,500
FEC30 (eggs/g)	17,674	13,814	700	73,800
WB (no. worms)	4,204	1,664	180	6,320
FL (mm)	19.31	1.67	15.39	23.71
FF (no.eggs)	422	163	131	816

Key: FECX, Fecal Egg Count at X dpi, WB, worm burden, FL, Average lengths of 35 sampled female worms/lamb, FF, average number of eggs recovered *in utero* of 20 recovered female worms/lamb

During hematological sampling, platelet agregates were detected for one individual that was not included any more in the analysis of hematological data. Average hct_0 values were in the common range of values (ranging from 29 to 46) as well as white cell counts (table 7.3). After experimental challenge, BCxBC lambs hematocrit decreased by 6.8% until 27 dpi, one individual achieving a drop of 15 points in its hematocrit (table 7.3). Blood loss was associated to a strong increase in reticulocytes cells as well as a platelet production as soon as 14 dpi (table 7.3). Considering white blood cells, no major modifications in numbers of circulating lymphocytes or granulocytes could be observed except a slight reduction after 27 dpi (see table 7.3).

7.3.2 First validation step: working on the selection sweep

The initial part of the experiment was based on the linkage analysis results that indicated two overlapping QTL affecting FEC in naive and infected lambs. In addition, an haplotype of 5 SNP under selection in the MBB breed was considered as contributing to the QTL likelihood. Lambs were produced and clustered into allelic groups to investigate the functional properties of this region under selection.

Results of the comparison between sweep-based allelic groups is provided in table 7.4. Note-

Table 7.3: Basic statistics of hematological parameters measured on BCxBC lambs

Variable	**Mean**	**Std**	**Minimum**	**Maximum**
hct0 (%)	35.87	3.21	29.00	46.90
hct14 (%)	33.27	3.50	26.40	41.20
hct27 (%)	29.19	3.63	22.60	38.80
mchc0 (g/dL)	33.88	1.53	29.80	36.60
mchc14 (g/dL)	33.78	1.61	30.70	39.60
mchc27 (g/dL)	31.30	1.87	27.00	34.60
lymph0 ($10^3/\mu$L)	5.90	1.26	3.58	9.23
lymph14 ($10^3/\mu$L)	5.40	1.15	3.28	8.68
lymph27 ($10^3/\mu$L)	4.28	1.28	1.66	7.91
mono0 ($10^3/\mu$L)	0.71	0.33	0.21	2.00
mono14 ($10^3/\mu$L)	0.70	0.28	0.23	1.60
mono27 ($10^3/\mu$L)	0.71	0.28	0.16	1.51
eo0 ($10^3/\mu$L)	0.11	0.11	0.02	0.68
eo14 ($10^3/\mu$L)	0.13	0.08	0.03	0.33
eo27 ($10^3/\mu$L)	0.06	0.05	0.01	0.31
baso0 ($10^3/\mu$L)	0.07	0.04	0.01	0.22
baso14 ($10^3/\mu$L)	0.05	0.04	0.02	0.20
baso27 ($10^3/\mu$L)	0.05	0.03	0.01	0.14
neut0 ($10^3/\mu$L)	2.71	0.76	1.44	4.38
neut14 ($10^3/\mu$L)	2.59	0.75	1.59	5.14
neut27 ($10^3/\mu$L)	2.34	0.76	1.30	5.02
ret0 (%)	0.06	0.03	0.01	0.13
ret14 (%)	0.05	0.02	0.01	0.10
ret27 (%)	0.43	0.72	0.01	3.42
plt0 ($10^3/\mu$L)	382.93	94.28	231.00	670.00
plt14 ($10^3/\mu$L)	443.47	96.73	185.00	656.00
plt27 ($10^3/\mu$L)	467.42	97.21	238.00	704.00

Key: hct_X, hematocrit at X dpi, mchc_X, hematocrit at X dpi, lymph_X, circulating lymphocytes at X dpi, mono_X, circulating, monocytes at X dpi, eo_X, circulating eosinophils at X dpi, baso_X, circulating basophils at X dpi, neut_X, circulating neutrophils at X dpi, ret_X, circulating reticulocytes at X dpi, plt_X, circulating reticulocytes at X dpi

worthy, no significant differences could be observed for FEC traits, nor for WB, even if slightly less eggs were excreted by the SW_{MBB} group. This was especially true in the early phase of infection (table 7.4), the difference between the SW_{MBB} and SW_{RMN} groups being of 0.59 phenotypic standard deviation (σ_p).

Circulating reticulocytes dramatically increased at 27 dpi in the SW_{RMN} group (p=0.02, table 7.4) which was consistent with hematocrit losses (p=0.06). In addition, basophils count at 27 dpi dropped in the same group (p=0.02, table 7.4).

Table 7.4: Comparison of mean parasitological and hematological traits of the sweep-based allelic groups

Trait[a]	**SW**$^{b}_{MBB}$	**SW**$_U$	**SW**$_{RMN}$	**Overall p-value**[c]	**Estimated difference**[d] **betw. SW**$_{MBB}$ **and SW**$_{RMN}$
FEC18	172 (373)	93 (209)	814 (1,204)	0.09	-0.59
FEC21	1,320 (1,571)	1,378 (2,048)	3,847 (4,001)	0.32	-0.36
FEC24	4,277 (2,803)	4,328 (3,277)	6,540 (5,772)	0.74	-0.27
FEC27	5,827 (3,244)	6,725 (4,671)	8,463 (6,993)	0.65	-0.23
FEC30	14,759 (7,340)	14,594 (10,282)	23,507 (19,077)	0.27	-0.35
WB	4,765 (1,172)	3,887 (1,850)	4,173 (1,729)	0.70	0.31
FL	18.88 (0.88)	18.90 (1.75)	20.10 (1.80)	0.06	-0.88
FF	376 (110)	374 (155)	515 (174)	0.02	-0.96
hct_{14}	33.2 (2.93)	34.7 (3.27)	31.8 (3.68)	0.06	0.50
hct_{27}	29.7 (2.67)	30.2 (3.94)	27.7 (3.59)	0.16	0.60
$baso_{14}$	0.06 (0.04)	0.06 (0.05)	0.04 (0.02)	0.95	-0.09
$baso_{27}$	0.06 (0.03)	0.07 (0.03)	0.03 (0.02)	0.02	0.36
ret_{14}	0.05 (0.02)	0.05 (0.03)	0.05 (0.01)	0.84	0.22
ret_{27}	0.12 (0.13)	0.25 (0.28)	0.85 (1.06)	0.02	-0.90

a: FECX: Fecal Egg Count at X dpi (in eggs/g of feces); WB: worm burden (no. worm/sheep); FL: female worm length (in mm); FF: female worm fertility (no.eggs/female worm); hct_X: hematocrit value at X dpi (in %); $baso_X$: circulating basophils count at X dpi; ret_X: circulating reticulocytes count; *b*:SW_{MBB}, homozygous MBB/MBB at the sweep region, SW_{RMN}, carriers of two RMN sweep allele carried by F1 sires contributing to the QTL, SW_U, carrier of only one MBB sweep allele. *c*: estimated on corrected data;*d*: estimated on corrected data and given in phenotypic standard deviation; *e*: average mean of raw phenotypes, with standard deviation in bracket

Contrasting these minor differences, female worms fertility significantly differed between the considered groups. Females were shorter of 1.5 mm in the SW_{MBB} lambs in comparison to the SW_{RMN} group (p=0.03) hence suggesting a putative sweep effect on worms fecundity (p=0.06, table 7.4). The gradual increase in average female fecundity measured between the SW_{MBB}, SW_U and SW_{RMN} groups reinforced this finding. On average, female worms recovered in SW_{RMN} lambs carried 156 eggs more than those sampled from the SW_{MBB} group, thus

representing a 0.96 σ_p difference (table 7.4).

Overall, no significant differences were found between SW_{MBB} and SW_U (data not shown), suggesting the MBB sweep allele exerts the same effect in homozygote and heterozygote individuals.

7.3.3 Estimated effect of the 4-SNP haplotype region

The BCxBC flock was created using the results of a within-family analysis performed on a BC population of 1,000 individuals and the selection sweep detected in the MBB breed. After the BCxBC were clustered into experimental groups according to their sweep genotype, an association analysis was performed on the BC data. At each position of the genome, the association analysis estimated an across-families effect of every haplotype segregating in the BC founders [454]. The maximal probability of a segregating QTL was mapped at 56.06 Mbp. Every haplotype identified in the BC founders at this position are listed in table 7.5.

Table 7.5: Estimated 4-SNP haplotype effect in the BC population

4-SNP allele[a]	**Frequency**	**Effect**[b]	**s.e.**[c]	**Note**
$AGCA_{MBB}$	0.25	-1.06	0.13	
$GAAG_{MBB}$	0.50	-0.71	0.13	Resistant alleles observed
$GGCA_{RMN}$	0.10	-0.69	0.13	in the BCxBC progenies
$GGCG_{MBB}$	0.25	-0.58	0.13	
$GACG_{RMN}$	0.19	-0.54	0.13	
$GGCG_{RMN}$	0.15	-0.49	0.13	
$GAAG_{RMN}$	0.15	-0.48	0.13	
$GACA_{RMN}$	0.02	-0.47	0.14	
$GAAA_{RMN}$	0.12	-0.47	0.13	Susceptible alleles segregating
$AGCG_{RMN}$	0.14	-0.43	0.13	in the BCxBC population
$AGCA_{RMN}$	0.09	-0.39	0.13	
$GGAG_{RMN}$	0.02	-0.21	0.15	Alleles specifically observed
$AAAG_{RMN}$	0.01	0.12	0.15	in the BC population

a: MBB, Martinik Black belly breed, RMN, Romane breed; *b*: given in phenotypic standard deviation; *c*: standard error

A first haplotype originating from the MBB breed ($AGCA_{MBB}$) was associated to the most favorable effect, *i.e.* -1.06 phenotypic standard deviation (table 7.5). Among most favorable haplotypes, no significant differences in the effect of the $GAAG_{MBB}$ and $GGCA_{RMN}$ haplotypes were observed (p-value=0.3). This $GGCA_{RMN}$ haplotype broke the MBB haplotypes hegemony among the most favorable haplotypes.

Indeed, the MBB allele had significantly more favorable effect than their respective IBS RMN version (p-values $< 10^{-4}$ in each case). This trend was particularly obvious for the AGCA$_{MBB}$ and GAAG$_{MBB}$ haplotypes, which were ranked as the two most favorable alleles whereas their RMN counterparts were in the second half of the ranking (table 7.5).

In addition, the most unfavorable effects were associated to the GGAG$_{RMN}$ and AAAG$_{RMN}$ haplotypes.

7.3.4 BCxBC clustering based on the 4-SNP alleles

Exploiting the results of the association analysis, the BCxBC lambs flock was clustered again according to the 4-SNP alleles they carried (table 7.1).

Due to the initial focus on the sweep region, some of the 4-SNP alleles segregating in the BC population and few homozygote individuals were swept off in the BCxBC population (see table 7.6). Therefore several alleles with *a priori* similar effects on FEC were used for tagging BCxBC as "resistant" or "susceptible". The group of lambs predicted to be resistant was denoted 4SNP$_R$ and its opposite 4SNP$_S$. To be included into one of these categories, every lamb belonging to one cluster did not carry any other allele selected for defining the group of opposite susceptibility. Remaining animals, *i.e.* carriers of "neutral" alleles or carriers of two alleles defining opposite susceptibility, were gathered into a 4SNP$_U$ group.

The 4SNP$_R$ group gathered individuals that inherited the most favorable allele (AGCA$_{MBB}$). Due to the limited size of this category (n=3), lambs carrying GAAG$_{MBB}$ or GGCA$_{RMN}$ were also added to this group. In the end, 16 lambs were in the 4SNP$_R$ group.

The opposite cluster consisted in predicted susceptible lambs (denoted 4SNP$_S$). None of the two most unfavorable alleles identified in the BC population (AAAG$_{RMN}$ and GGAG$_{RMN}$) segregated in the BCxBC population any more. This was in relationship with their original low allelic frequencies in the BC population founders, *i.e.* 1 and 2% respectively. Hence, every individual carrying the AGCG$_{RMN}$ allele or the AGCA$_{RMN}$ allele but none of the three most favorable alleles were considered for this group (n=5 in total). To increase the $4SNP_S$ group sample size, seven animals carrying GACA$_{RMN}$ or GAAA$_{RMN}$ alleles were also added.

The effect of the rest of the genome estimated by the average genomic values was similar between 4SNP groups (p-value=0.37, table 7.1).

Table 7.6: Observed 4-SNP based genotypes in the BCxBC population and associated frequencies

Allelic *group*[a]	**Allele 1**	**Allele2**	**Frequency**
	$GGCG_{RMN}$	$GAAG_{MBB}$	2
	$GGCG_{RMN}$	$AGCA_{MBB}$	1
	$GGCG_{MBB}$	$GGCA_{RMN}$	1
	$GGCG_{MBB}$	$GAAG_{MBB}$	2
	$GAAG_{MBB}$	$GGCA_{RMN}$	1
$4SNP_R$	$GAAG_{MBB}$	$GACG_{RMN}$	4
(n=16)	$GAAG_{MBB}$	$GGCG_{MBB}$	1
	$GAAG_{MBB}$	$GAAG_{MBB}$	2
	$GAAG_{MBB}$	$AGCA_{MBB}$	1
	$AGCA_{MBB}$	$GGCG_{MBB}$	1
	$GGCG_{RMN}$	$AGCG_{RMN}$	1
	$GGCG_{RMN}$	$GAAA_{RMN}$	2
	$AGCA_{RMN}$	$AGCA_{RMN}$	1
$4SNP_S$	$GACA_{RMN}$	$GGCG_{RMN}$	4
(n=12)	$GACA_{RMN}$	$GACG_{RMN}$	1
	$GACA_{RMN}$	$GAAA_{RMN}$	2
	$GACA_{RMN}$	$GGCG_{MBB}$	1
	$GGCG_{RMN}$	$GGCG_{RMN}$	4
	$GGCG_{RMN}$	$GACG_{RMN}$	1
	$AGCA_{RMN}$	$GAAG_{MBB}$	1
	$GACA_{RMN}$	$GGCA_{RMN}$	1
	$GGCG_{MBB}$	$GGCG_{RMN}$	2
$4SNP_U$	$GGCG_{MBB}$	$GACG_{RMN}$	2
(n= 16)	$GGCG_{MBB}$	$GGCG_{MBB}$	1
	$GAAG_{MBB}$	$AGCA_{RMN}$	1
	$GAAG_{MBB}$	$GACA_{RMN}$	1
	$GAAG_{MBB}$	$AGCG_{RMN}$	1
	$GAAG_{MBB}$	$GAAA_{RMN}$	1

a:$4SNP_R$, predicted resistant, $4SNP_S$, predicted susceptible, $4SNP_U$, unknown status

7.3.5 Phenotypic comparison of the 4-SNP-based BCxBC categories

Results from the comparison between each of the three susceptibility groups, *i.e.* 4SNP$_R$, 4SNP$_S$ and 4SNP$_U$, are provided in table 7.7.

Table 7.7: Comparison of mean parasitological and hematological traits of the 4SNP-based allelic groups

Trait[a]	**4SNP**$_R$[b]	**4SNP**$_U$	**4SNP**$_S$	**Overall p-value**[c]	**Estimated difference**[d] **betw. 4SNP**$_R$ **and 4SNP**$_S$
FEC18	100 (221)[e]	169 (349)	955 (1305)	0.05	-0.88
FEC21	1,205 (1,776)	1,377 (2,203)	4,642 (3,850)	0.05	-0.59
FEC24	4,319 (2,810)	3,613 (3,469)	8,013 (5,485)	0.16	-0.43
FEC27	6,169 (3,612)	5,469 (4,384)	10,492 (6,944)	0.23	-0.27
FEC30	13,638 (7,452)	13,166 (8,308)	29,067 (19,526)	0.06	-0.65
WB	4,469 (1,441)	3,798 (1,921)	4,392 (1,607)	0.64	0.24
FL	18.9 (1.1)	18.7 (1.7)	20.6 (1.7)	0.004	-1.02
FF	353 (105)	360 (112)	598 (162)	< 0.0001	-1.50
hct$_{14}$	34.4 (3.5)	33.3 (3.5)	31.8 (3.3)	0.06	0.66
hct$_{27}$	30.7 (3.5)	29.3 (3.4)	27.1 (3.4)	0.03	0.86
baso$_{14}$	0.06 (0.04)	0.06 (0.05)	0.04 (0.02)	0.26	-0.25
baso$_{27}$	0.07 (0.03)	0.06 (0.03)	0.03 (0.02)	0.04	0.67
ret$_{14}$	0.05 (0.03)	0.05 (0.02)	0.05 (0.01)	0.63	0.00
ret$_{27}$	0.11 (0.10)	0.36 (0.63)	0.91 (1.00)	0.03	-0.93

a: FECX: Fecal Egg Count at X dpi (in eggs/g of feces); WB: worm burden (no. worm/sheep); FL: female worm length (in mm); FF: female worm fertility (no.eggs/female worm); hct$_X$: hematocrit value at X dpi (in %); baso$_X$: circulating basophils count at X dpi; ret$_X$: circulating reticulocytes count; *b*:4SNP$_R$, predicted resistant, 4SNP$_S$, predicted susceptible, 4SNP$_U$, unknown status. *c*: estimated on corrected data;*d*: estimated on corrected data and given in phenotypic standard deviation; *e*: average mean of raw phenotypes, with standard deviation in bracket

As for the sweep-based groups, the 4SNP region identified individuals with strong differences in female worms length and fertility, a 1 σ_p contrasting phenotypes of the 4SNP$_R$ and 4SNP$_S$ lambs 7.7. In addition, the 4-SNP-based selection resulted in picking up true "high-" and "low-FEC" animals as illustrated by the 0.88 and 0.65 σ_p obtained between the 4SNP$_R$ and 4SNP$_S$ lambs at 18 and 21 dpi (p=0.05 in both cases,7.7). This trend disappeared on the following FEC sampling but almost reached significance for FEC30 (p=0.06, table 7.7). However no significant differences was observed for WB.

Further, 4SNP$_R$ lambs also exhibited significantly less blood loss at 27 dpi (p=0.03) and to a lesser extent at 14 dpi (p=0.06). This was concomittant of a higher production of reticulocytes in the 4SNP$_S$ lambs. These reduced blood losses have already been outlined in the SW$_{MBB}$

group.

No significant differences between the $4SNP_R$ and $4SNP_U$ animals could be found for any of the considered traits, as already observed with the sweep-based groups. The $4SNP_U$ group was composed of equal amount of lambs carrying "neutral" RMN alleles (n=5), lambs carrying two alleles with opposite effects (n=5) or lambs carrying the third MBB allele and one unfavorable allele (n=6). This clustering of genotypes and the associated phenotypes would be in favor of a dominance effect of the favorable alleles over the others.

7.3.6 Testing for differential candidate gene expression between the selected groups

A gene expression comparison was performed between carriers of the $GAAG_{MBB}$ allele and carriers of the $GAAG_{RMN}$ allele. These two alleles were chosen as being the most frequent in the sampled $4SNP_R$ and $4SNP_S$ groups. These two allelic group used for gene expression comparison were denoted as R_I (n=9) and S_I (n=8) respectively.

Table 7.8: Fold change in gene expression according to the tissue sample and compared groups

		Comparison performed			
Tissue	**Tested gene**	$\mathbf{R}_I$ / $\mathbf{S}_I^{a,b}$	$\mathbf{R}_I$ / $\mathbf{R}_C$	$\mathbf{S}_I$ / $\mathbf{S}_C$	$\mathbf{R}_C$ / $\mathbf{R}_C$
	GAL15	7.02	575.63*	82.03	0.60
	ITLN2	6.97	22.44*	3.22*	4.94
	TFF3	1.61	6.86*	4.26	1.50
Fundus	IL4	4.07*	2.22	0.54	1.16
	IL13	4.51*	4.79*	1.06	0.93
	IFNγ	0.83	1.13	1.36	1.47
	TNFα	0.85	0.63	0.74	1.23
	IL4	1.22	2.33*	1.91	0.81
	IL13	1.24	0.50	0.41	0.78
Lymph node	IFNγ	0.69	0.33	0.48*	1.08
	TNFα	0.70*	0.62	0.88	1.12
	CCL16	1.65	1.13	0.69	0.87
	OX40	0.68	0.59	0.87	1.11

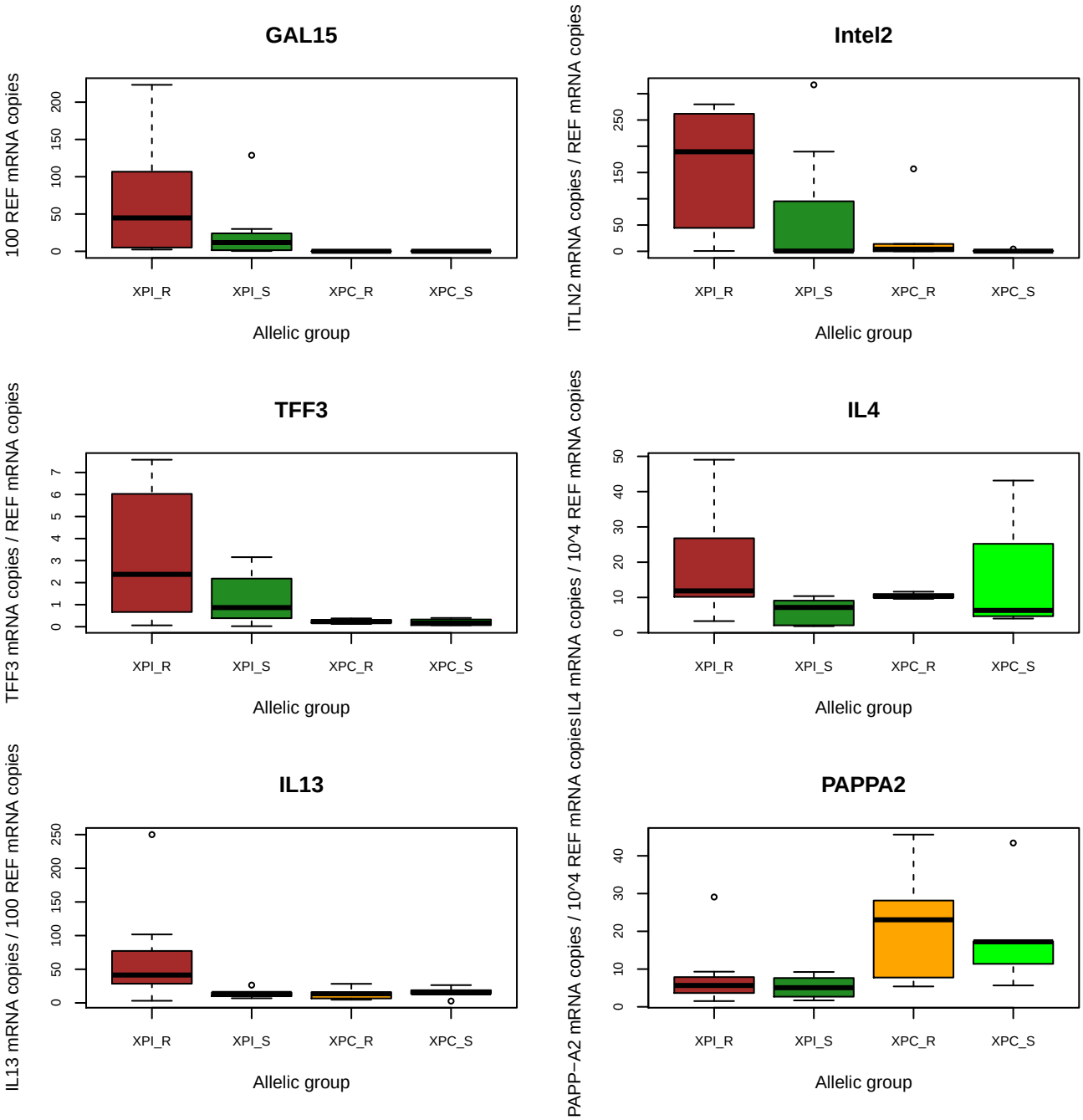
a:R_I, infected predicted resistant, S_I, infected predicted susceptible, R_C,control predicted resistant, S_C, control predicted susceptible; *b*: * indicates significant differences ($p < 0.05$)

Control individuals (denoted R_C and S_C for predicted resistant and susceptible respectively) were picked up among the 17 unchallenged lambs. Provided selection of this control animals had been done before the association analysis was done, no previous control of their 4-SNP QTL allele had been done. Hence, only two individuals were available for each category, *i.e.* R_C and

S_C. To obtain better estimates of basal level of gene expression in resistant and susceptible animals, three individuals carrying another resistance/susceptibility allele were added to each group.

In total 27 animals were retained for gene expression measure whose results are presented in table 7.8 and figure 7.2.

Figure 7.2: Gene expression levels in BCxBC abomasal mucosa



Lectin genes (*GAL15* and *ITLN2*) and the *TFF3* gene showed the highest induction after

infection (table 7.8 and figure 7.2). The over-expression of *GAL15* and *TFF3* mediated by *H. contortus* was significant between the R_I and the R_C individuals (p=0.04 and 0.02, figure 7.2). However, no significant differences was found between infected groups for these innate-response-related genes.

Among the tested cytokine genes, a 4-fold over-expression of the IL-4 and IL-13 genes was measured in abomasal mucosa of the R_I group in comparison to the S_I sheep (p=0.04 and p=0.02 for IL4 and IL13 respectively, table 7.8). None of these cytokines were differentially expressed in the draining lymph node. Still, IFN_γ expression in lymph node was significantly reduced by a 1/3 factor in the R_I in comparison to the basal level observed in the R_C (p=0.01, table7.8), suggesting a down-regulation of the Th-1 biased immunity after nematode infection. Noteworthy, an almost significant down-regulation of the OX40 (p=0.07) was also observed in lymph node of the R_I group compared to the S_I lambs.

The slight over-expression of the TNF_α cytokine in the S_I sheep's lymph nodes was the only additional differential expression found between R_I and S_I lambs (table 7.8).

7.3.7 Looking for positional candidate genes

Table 7.9: Annotated genes lying between 55.1 and 57.1 Mbp on the sheep genome

Gene	**Starting position**[a]	**Related functions**
RFWD2	55,199,014	DNA damage response
SGOL1	55,509,260	Cell cycle
PAPP2	55,757,260	Regulation of Insulin-like Growth Factor (IGF) activity
ASTN1	56,095,325	Nervous system development
FAM5B	56,545,767	Cell cycle , nervous system development
TAF9	56,864,943	Initiation of transcription by RNA polymerase II

Retrieved from the http://www.livestockgenomics.csiro.au/cgi-bin/gbrowse/oarv3.0/ website, last accessed on the 5 september 2012. a: position in bp.

In addition to the genes with previous knowledge of differential expression between pure breeds, annotated genes of the sheep genome were retrieved within a 2-Mbp region centered at 56.1 Mbp. The six genes lying in this chromosomal segment are listed in table 7.9.

Among these genes, the *ASTN1* (Astrotactin-1) and the *PAPP-A2* (Pregnancy-associated plasma protein-A2 gene) loci laid just below the 56.06 Mbp position. For time purpose, the *ASTN1* gene whose functions in neurological development could not be directly related to resistance to GIN infection was not investigated. The *PAPP-A2* has been involved in fetal develop-

ment. It is hence linked to the Insulin Growth Factor (IGF) gene which is known to participate in the anti-helminth immune response [89]. A potential relationship was hypothesized and the *PAPP-A2* expression was measured. However every considered group showed similar expression level either before or after infection in each of the considered tissues (figure 7.2).

7.4 Discussion

7.4.1 First published functional investigation of a QTL affecting resistance to nematodes in sheep

The reported study aimed at investigating the functional properties of a QTL region associated to resistance to GIN. Abundant literature has been produced on the role of functional candidate genes like *MHC* [407, 458, 488] or IFN_γ [94, 119]. Other research teams mined the functional differences between divergent lines of sheep selected for low or high FEC [201, 245, 244]. But to our knowledge, this is the first time a functional investigation of a QTL affecting resistance to GIN in sheep is reported.

Two goals were pursued while producing BCxBC animals. The first objective was to validate and dissect the architecture of the identified QTL through the comparison of individuals selected on their particular QTL alleles. For this comparison, two regions were investigated: either the selection signature identified in the MBB breed or the 4-SNP haplotype that maximized the likelihood in the association analysis. A second related goal was to precise the functional mechanisms of the QTL, thanks to a wider range of phenotypes, including gene expression analysis.

7.4.2 What validation has been achieved ?

The original QTL had been found in a familial experiment, two families contributing the most to the likelihood. In their guidelines, Lander & Kruglyak indicated that for validation, a QTL should be detected in an independent population [291]. This kind of experiment is a huge work load and it is also highly subjected to failure. Before getting outside the experimental populations, it was decided to produce progenies selected to carry two favorable/unfavorable QTL allele.

During the BCxBC creation process, a limited number of BC sires were selected according to the particular QTL alleles they inherited from their F1 sires. The 71203 sire was an outstanding

contributor to the predicted susceptible groups (14 out of the 15 lambs) as the other selected sire for this susceptibility group was not able to mate selected females. Therefore, comparing predicted susceptible animals based on the OAR12 QTL to the other allelic groups was similar to comparing one family to others. In the end, the QTL effect should be confounded by the average 71203 sire's breeding value. To take care of this point, gEBV were estimated and fitted as covariate while testing for any significant difference between the QTL allelic groups. This was thought to correct for the effect of the rest of the genome. In addition, no significant differences could be observed between compared groups.

7.4.3 The sweep region was not predictive of the observed FEC

The sweep genotype was taken as an indicator for allelic class assignment of BCxBC lambs. Indeed, a selection signal was found in the MBB breed but still segregated in the RMN breed. Further, the fitting of the 5-SNP-swept haplotype as a fixed effect in the QTL detection model eroded the QTL profile (see 6.3). It was hence hypothesized that the sweep might have been mediated by high GIN infection pressure and might contribute to the QTL likelihood.

After experimental infection, sweep-based groups showed differences in female worm fecundity (0.96 σ_p between SW_{MBB} and SW_{RMN}) and in hematocrit drop (higher blood loss at 27 dpi SW_{RMN}, assessed by significant increase in the ret_{27} count). However, the lack of difference in eggs excretion suggests the sweep region was sub-optimal for the identification of true "resistant" or "susceptible" lambs. Our working hypothesis was rather strong as the selection sweep detection only relied on population genetics data without phenotypic support. In addition, sweep signal and maximal LRT position for FEC at first infection were 15 Mbp far apart. Still, this region was efficient at discriminating between individuals able to control the female worms fertility. Whether this reflects a true sweep-associated property or a simple association between the sweep and the causative mutation remains undetermined.

7.4.4 A 4-SNP haplotype tags true resistant and susceptible lambs

Focusing on the 4-SNP haplotype was more efficient at discriminating between BCxBC lambs. For instance, the sweep genotype explained 17 and 12% of the observed variation for FF and FL respectively. The 4-SNP-based tagging increased the explained part of observed variation to 44 and 24% of these traits respectively. Extreme genotypic groups built with this indicator not only highlighted strong differences in worm fecundity and blood loss, but also pointed significant

variation in FEC (-0.8 and -0.43 σ_p between 4SNP$_R$ and 4SNP$_S$ groups). It is worth noting that the obtained difference for FEC21 were those expected from the haplotype effects estimated in the BC population, *i.e.* a -0.38 σ_p deviation between the two groups of haplotypes considered. The over-estimation for FEC18 might lie in the fact that many 0 values were obtained for the 4SNP$_R$ group (10 out of 16 observations) that may artificially shrink average 4SNP$_R$ FEC towards lower values.

7.4.5 A first step in dissecting biological properties of the QTL

Proposing a detailed mode of action of the investigated QTL region requires more investigation. We reported herein a QTL region whose most extreme alleles greatly affect female worm fecundity. This finding has also been reinforced by the expression data analysis that highlighted a difference in the IL4 and IL13 cytokines that contribute to the mounting of a Th-2 type environment. So far as we know, a QTL affecting *in utero* eggs count has never been involved in a QTL detection study and it hence provides an original finding in the landscape of QTL mapping.

No density-dependent regulation of female worm fecundity [158, 100, 484] was observed in our data as every group showed similar worm burden. Further, every lamb was inoculated with the same infection dose. Therefore, the strong differences in female fertility reported here cannot be related to some competition for food resources eventually occuring in heavily infected animals, nor to the known compensatory reduction in fecundity associated to heavy infection load [158, 484]. Hence, observed differences in worm fecundity seem to be directly mediated by the QTL under study. No other factor had a significant effect on this trait. However, the QTL allelic group explained 40 % of the variation at most, suggesting other factors mediate this trait as reported elsewhere [158, 484].

Further, significant differences in hematocrit were also consistently observed between carriers of the favorable and unfavorable alleles (illustrated by reticulocytes count in sweep group). Provided reticulocytes production was significantly higher in susceptible lambs after infection only, observed differences translated true blood loss differences and not a better regeneration ability of the resistant lambs. In addition, hct$_{27}$ was negatively correlated to FL (-0.43) and FF (-0.53). Both findings lead to the hypothesis that the investigated QTL region could limit worm feeding hence reducing their growth (shorter females) and fecundity (lower eggs recovered *in utero*).

From the expression data available on lectins and TFF3, it seems that the QTL region does

not affect the tested innate components of the rejection response. On the contrary, it may rather act on the acquired response, as lambs carrying one MBB favorable allele demonstrated higher amount of Th-2 cytokines expression in abomasal mucosa. Further, the IFN_{γ} expression was also down-regulated in R_I in comparison to the unchallenged lambs of the same susceptibility group. The additional slight down-regulation of OX40 between the same two groups constitutes another factor contributing to the mounting of a Th-2 environment against GIN while repressing the Th-1 response. Indeed it has demonstrated that the inhibition of the OX40 cytokine resulted in a more efficient expulsion of helminth in two mice models [243, 140]. Same findings were reported by Terefe *et al.* (2007) who compared pure breeds cytokines gene expression at 4 and 30 dpi [507] and Liénard *et al.* (2011) who performed a micro-array experiment at 8 dpi [316].

Simultaneous findings on female fecundity and Th-2 response are in good agreement with the existing literature. Indeed, Stear *et al.* proposed that local IgA response limits worm fecundity in *T. circumcincta* [484, 483] and during *H. contortus* infection [288]. More recent findings also support the relationship between the CD4+ T-cells number and the female length [445, 191]. This QTL region may thus act on the CD4+ T cells activation. Additional investigations on the characterization of subpopulations of T-cells in allelic carriers of each type could bring additional insights. As well, histological examination of abomasal mucosa of extreme animals could also confirmed the stronger Th-2 response by measuring the eosinophils infiltration and the number of mast cells.

These findings are interesting in two different ways. Firstly, it shows that the selection of lambs based on a 4SNP region resulted in a differential cytokinic environment in abomasal mucosa at 30 dpi. Even if R_I individuals did not carry any unfavorable allele and conversely, it seems that the 4 SNPs picked-up the differences observed in pure breed animals. Indeed genomic breeding values were similar between the two considered groups, so that average effect of the rest of the genome should have been the same between groups. Secondly, the two alleles chosen for gene expression study were IBS and only differed by their breed origin. This functional difference supports the estimated haplotypes effects in the BC population that showed significant differences between IBS alleles coming from the MBB and the RMN breed.

However, underlying functional genes contributing to the QTL are still missing. One of the nearest genes, *i.e.* the *PAPP-A2* gene, encodes a protease that cleaves the Insulin Growth Factor [559] which is involved in the immune response, and to bind with lectins that are involved in GIN recognition [528]. Still, this gene has mostly been associated to pregnancy and fetal development,

and there is no wonder in the lack of significant variation in the tested groups. The *ASTN-1* expression could have been tested as well as it perfectly matches the 4-SNP haplotype. Still, its associated function, *i.e.* neurological development [557], seems to be totally independent of the GIN infection. Other candidates are not too far apart like tenascin genes (TNN and TNR). The tenascin C gene has been involved in immune recognition of bacteria [309] and an eventual role in helminth infection cannot be ruled out.

7.4.6 What further investigation for dissecting the two investigated QTL ?

7.4.6.1 Meta-analysis

One way to increase the precision and power of QTL mapping is to examine other existing datasets and to perform meta-analysis [264, 413, **?**]. Though, heterogeneity between datasets (*e.g.* various environmental factors, differential linkage disequilibrium across populations, phenotypic heterogeneity) requires an appropriate correction [413]. Such a meta-analysis is currently under completion in the frame of the EU-funded 3-SR project (S.C. Bishop, personal communication). In this case, datasets are strongly heterogeneous: several breeds are considered and sheep were either exposed to natural challenge (mix of worms species with predominant *T. circumcincta*) or to experimental infection by *H. contortus*. Although it has been suggested that genetic resistance to one species holds for another species [199] and that experimental challenge and natural challenge triggered same genetic mechanisms [201], these variations may complicate analyses.

7.4.6.2 Developing new markers

As it has already been discussed before (9.2), the number of observed recombinations conditions the upper limit of the precision of mapping. Increasing marker density can also helps reducing the QTL confidence interval.

In the frame of this experimental population, developing new SNPs might add a bit more of information. It might be worth to develop new micro-satellites markers. These markers are more polymorphic and they could help defining new subset of progenies hence being more efficient at tagging QTL allele carriers. However, the information they bring is different as they usually belong to non-coding sequences.

7.4.6.3 Exploiting technology advances

The Genomics era has almost already come to an end and we are entering the "post-genomic" era. Indeed, sequencing a region of more than 1 Kb by sequence capture has now become more expensive than re-sequencing the whole genome. In addition, this re-sequencing brings additional data for improving the ovine genome assembly.

Obtaining sequences of the QTL region in groups of animals with opposite QTL alleles and a common ancestor would help determining new SNPs for tagging the true QTL location. Still, the additional value of this approach is not straightforward. SNPs have increased the marker density on OAR12 by a factor 10. However, progress in terms of precision has been only limited to a few Mbp. Even if the functional validation points out a 4-SNP haplotype region, the QTL region we need to consider is still large because of the LD extent in the BC population. Therefore, choosing resistant and susceptible animals on the basis of their QTL genotype still remains difficult. In this case, it thus seems that whole genome re-sequencing would only be a technological headlong rush.

On the contrary, RNAseq approach would simultaneously give access to the functional candidate and to its location [548, 404]. LD extent in the population would not interfere as this technique measures the relative abundance of transcripts. Any differential expression in one allelic group could be either:

1. the gene underlying the QTL
2. a gene regulated by the QTL region
3. a false-positive

In the first case, mapping the candidate sequence to the QTL confidence interval would be a strong evidence of causality. The two other options are more difficult to handle. After sampling the whole set of differentially expressed genes, gene network analysis might help understanding different interactions between regions of the genome and seeing any gene regulation. The third case will be highly dependent of technical issues (heterogeneity in tissue sampling or in sample processing for RNA extraction, RNA sequencing ...), of the sequencing depth and the relative abundance of transcripts, and of the considered sample size.

7.5 Conclusion

Our study aimed at investigating the functional properties of a QTL region affecting GIN infection in sheep. BCxBC lambs were selected based on their genotype at a selection sweep detected within the QTL in a MBB breed. A first attempt to compare the performance of individuals selected on this sweep region demonstrated significant impacts on female worms fertility and on blood losses (assessed by reticulocytes count at 27 dpi). Benefiting from the results of an association study conducted in the BC population, BCxBC lambs were clustered again into groups based on their allele at a 4-SNP region. The comparison of 4-SNP groups phenotypes not only confirmed the findings obtained with the wide QTL region, but also resulted in differences in FEC output. A gene expression study comparing carriers of one MBB-inherited allele against carriers of the same IBS allele from the RMN breed, indicated a stronger Th-2 environment in the abomasal mucosa of MBB-allele carriers.

These findings reinforce the actual presence of a QTL on OAR12 in this experimental population. They are also in favor of a QTL region that affects *H. contortus* females fecundity, and also regulates the cytokinic environment in abomasal mucosa under GIN infection. These phenomena and the associated effect on host blood loss variation may be related.

However no functional candidate genes can be proposed so far. Two major research paths could now be explored. Firstly, a proper validation comparing homozygote sheep selected on the 4-SNP allele they carry should be performed. The most unfavorable alleles identified in the BC population that had been swept off during the BCxBC selection process, should also be included in this comparison.

Secondly, the validation of the QTL should be undertaken in other populations. This is a major step for its use for breeding purpose, as the QTL has been identified in a familial design. This could be assessed by a meta-analysis of other datasets. The identified alleles could also be looked for in one of the pure breed to investigate their effect on GIN infection. In case this QTL also segregates in the RMN breed, it could be interesting for the RMN breeders to increase the relatively low allelic frequency of the favorable RMN allele (10% in the BC founders) while getting rid of the most unfavorable alleles that still segregate at very low frequency ($< 10\%$). Another solution could be to introgress MBB allele in the RMN breed.

In case the findings we report can be validated, it would be interesting to resequence interesting candidates with opposite genotypes/phenotypes. This would both help finding underlying

genes while contributing to the ovine genome assembly.

Chapter 8

Candidate gene approach on a QTL affecting pepsinogen concentration (manuscript in preparation)

Résumé

La concentration en pepsinogène plasmatique est utilisée comme un indicateur d'infestation à l'échelle du troupeau. Cependant, les mécanismes précis à l'origine de l'augmentation de la concentration en pepsinogène durant une infestation par *H. contortus* sont encore méconnus. Notre étude visait à valider un QTL affectant la variation de concentration en pepsinogène durant une haemonchose expérimentale.

Deux groupes d'animaux croisés BCxBC ont été sélectionnés sur la base de l'allèle au QTL dont ils disposaient. Ces animaux ont ensuite été infestés par *H. contortus* et suivis durant un mois après infestation.

Des différences significatives de variation de concentration en pepsinogène à J15 ont été observées entre les deux groups sélectionnés. Le séquençage de deux couples d'animaux génétiquement prédits "fort" et "bas" producteurs de pepsinogène a révélé 9 marqueurs introniques, non partagés entre les groupes. Le génotypage des autres animaux de chaque groupe permettra de confirmer l'éventuelle association de ces marqueurs à la concentration en pepsinogène.

Nos résultats ont également montré que la région QTL considérée affectait la résistance générale à *H. contortus*. De plus, les corrélations phénotypiques estimées montrent une relation favorable entre augmentation de la concentration en pepsinogène et l'intensité d'oeufs excrétés.

Dans l'état actuel des investigations, aucune explication ne peut être argumentée. Cependant, la proximité entre les gènes *CD5/CD6*, affectant la régulation des lymphocytes T et *PGA5*, qui code pour le pepsinogène, pourrait expliquer les tendances observées entre pepsinogène et résistance à *H. contortus*. Cette hypothèse devra être confirmée par, notamment, un séquençage supplémentaire des gènes *CD5* et *CD6* chez les animaux sélectionnés.

Summary

Serum pepsinogen has been considered as an indirect indicator for monitoring GIN infection at flock level. Still, global understanding of serum pepsinogen rise occurring during haemonchosis has been lacking. Our study aimed at validating a QTL affecting pepsinogen concentration variation during *H. contortus* infection.

Two groups of BCxBC animals were built based on the QTL allele they carried, before being infected by *H. contortus* and monitored for one month.

Significant differences were observed for pepsinogen variation at 15 dpi between selected groups, reinforcing findings of the QTL study. The most consistent candidate gene was the PGA5 locus that codes for pepsinogen. Sequencing of two couples of predicted "high-" and "low-responders" unraveled nine markers specific to each group and located in intronic sequences. Additional genotyping of remaining animals should confirm/reject these markers.

Our findings also suggested the QTL under investigation could affect the outcome of GIN infection, and favorable correlations were reported between pepsinogen concentration and eggs output. No obvious explanation can be proposed. Still, the close relationship between the *CD5/CD6* loci and the *PGA5* locus might explain colinearity between the QTL allele effect on serum pepsinogen and infection outcome. This hypothesis requires further investigation before being confirmed.

8.1 Introduction

Pepsinogen has been used as a marker of exposure to GINs that settle in abomasum, *i.e. H.contortus* and *T. circumcincta* in sheep [340, 507, 164, 472], and *Ostertagia ostertagi* in cattle [380, 87].

In a previous QTL analysis, 228 Martinik Black-belly x Romane back-cross lambs have been

measured for pepsinogen concentration before and 15 days after an experimental infection by *H.contortus* [454]. A narrow QTL region around 37 Mbp has been associated to the variation of pepsinogen concentration on OAR21. Within this region, the PGA5 locus that codes for the pepsinogen protein was identified as a consistent candidate gene. In addition, the same region had been linked to FEC at first infection, and favorable correlation (-0.21) had been estimated between 271 parasite eggs output and pepsinogen concentration at reinfection [454].

In order to better understand the genetic architecture underlying the variation of pepsinogen concentration during haemonchosis, we propose to validate the identified QTL region by investigating the *PGA5* candidate locus in a BCxBC flock.

8.2 Materials and methods

8.2.1 Animal materials

Following the QTL detection study, crosses between BC animals have been performed based on their genotype at another QTL region mapped on OAR12. Individuals from a first generation of BCxBC crosses were also available (see 8.1.2). After selection, two BC sires and one BCxBC sire were mated to 82 BC and BCxBC ewes of corresponding QTL genotype. In December 2010, 130 BCxBC lambs were born at the INRA La Sapinière experimental unit. In the end, 115 lambs were available for experimental infection.

8.2.2 Experimental design and measured traits

Among the 115 lambs available for infection, a first flock of 61 animals (**flock 1**) was challenged and intensively monitored for a one month period (see chapter 8, section 8.1.2).

The second flock (**flock 2**) was kept indoor at the INRA La Sapinière experimental unit from birth to the starting of the experimental infection, *i.e.* one week before the flock 1 infection. Flock 2 was also applied same drenching procedure but no anti-coccidiosis treatment was performed. At five months of age, flock2 lambs were given 10,000 *H. contortus* larvae from the same production batch as for flock1. At 30 dpi, lambs were drenched with ivermectin (ORAMEC, 2.5 mg/10 kg bodyweight, Merial, France). Feces were collected twice at 21 and 30 dpi for FEC determination (**FEC21** and **FEC30** respectively) following the same protocol as for flock 1. Three blood samples were collected (one before infection and two others at 14 and 30 dpi) for packed-cell volume determination (hct_0, hct_{14}, hct_{30} respectively).

In addition, serum pepsinogen concentration was also determined for every infected BCxBC animal following the micro-method of Dorny & Vercruysse [129]. Briefly, the serum sample was acidified with HCl and incubated overnight at 37C with bovine serum albumin (BSA) before being stopped with 4% trichloro-acetic acid (TCA). The resulting mixture was centrifuged at 14,000 rpm for 5 minutes. An aliquot from the supernatant was added to 0.25 M NaOH and the plates were incubated at room temperature with folin reagent for 30 minutes. The liberated tyrosine was estimated by reading the absorbance at 680 nm and the values were expressed as milliunit (mU) tyrosine/litre of serum. Three time points were considered: before infection ($\mathbf{peps}_0$), at 15 dpi ($\mathbf{peps}_{15}$) and at 30 dpi ($\mathbf{peps}_{30}$).

8.2.3 SNPs genotyping and editing

One month after birth, blood sample was taken and sent to LABOGENA (www.labogena.fr) for genotyping with the ovine 50K SNP chip (Illumina, San Diego, CA). The working list of SNPs described in Sallé et al. [454] was used. It was checked that none of these SNPs had a MAF below 1% and that their call rate was above 97%. Updated SNP positions from the third release of the ovine genome assembly were considered for analysis.

8.2.4 Clustering of animals according to their QTL allele

SNP positions were updated according to the latest sheep genome assembly released in june 2012. In addition, some changes were brought to the haplotype effect estimation implemented in the QTLMAP software. Taking these two updates into account, the within family analysis and LD-based analyses were repeated in the BC population on OAR21 as described in [454].

Following the QTL analyses, chromosomes of every lamb were reconstructed using the LinkPhase software [133], so that it was possible to determine the sequence of 4 markers they inherited at the QTL position. BCxBC lambs were clustered into two allelic groups. A "high-responder" (HR) group was defined as lambs carrying the haplotype associated to the highest pepsinogen concentration increase under infection. Conversely, a "low-responder" (LR) group consisted in lambs carrying the haplotype with the lowest effect on pepsinogen variation under infection. Individuals carrying both haplotypes or carrying none of them were not considered any more.

8.2.5 *PGA5* sequencing

Following the QTL mapping analysis, the CSIRO genome browser was used to screen for any candidate gene. The *PGA5* gene perfectly matched the maximal likelihood position and was an obvious candidate for further investigation as it encodes the pepsinogen. Two animals per allelic group were chosen for gene sequencing and for screening SNP markers. This work was undertaken by K. Tabet and F. Woloszyn under the direction of G. Tosser-Klopp at the Laboratoire de Génétique Cellulaire (INRA, Toulouse).

8.2.6 Statistical analyses and transformation applied to phenotypes

Normality was checked for every trait using the UNIVARIATE procedure implemented in SAS (SAS Institute Inc., SAS 9.1.3 Help and Documentation, Cary,NC: SAS Institute Inc., 2000-2004). FEC showed strong departure from normality and were thus applied a fourth root transformation.

The pepsinogen variation between $pepsi_{15}$ and $pepsi_0$ (dpeps1) and between $pepsi_{30}$ and $pepsi_0$ (dpeps2) were considered for analysis. Individuals that showed a decrease in pepsinogen concentration during infection were removed from the dataset. Outliers were discarded as well. Due to the incubation period inherent to the pepsinogen titration, samples could not be handled altogether the same day. The date of experiment was thus considered as a fixed effect to account for potential variations from one day to another.

Correlations were estimated using the CORR procedure from the SAS software, while correction for environmental fixed effects and testing for the QTL haplotypes effect was performed using the MIXED procedure (SAS Institute Inc., SAS 9.1.3 Help and Documentation, Cary, NC:SAS Institute Inc.,2000-2004). Flock and sex were considered as fixed effect. Basal value (measured before infection) was fitted as covariate when testing for any allelic effect on measurements under infection.

8.3 Results

8.3.1 Parasitological traits in the BCxBC flock (table 8.1)

Basic statistics of FEC and PCV traits are listed in table 8.1. FEC were significantly higher in flock1 ($p=0.02$ and $p<0.0001$ for FEC21 and FEC30 respectively) but hematocrit values were similar in both flocks after infection ($p= 0.49$ and 0.42 for hct_0, hct_{14} and hct_{30} respectively).

However hct0 values determined in flock 2 were slightly lower than in flock 1, which is certainly due to the measurement method (automatic determination in flock 1 and micro-hematocrit in flock 2).

Table 8.1: Comparison of FEC and hematocrit in the two BCxBC flocks

	flock 1					**flock 2**				
Trait	N	Mean	Std	Min.	Max.	N	Mean	Std	Min.	Max.
FEC21 (eggs/g)	44	2,205	2,978	0	12,500	58	1,753	3,227	0	14,750
FEC30 (eggs/g)	44	17,674	13,814	700	73,800	58	8,884	6,764	0	28,600
hct_0 (%)	44	35.16	5.66	4.80	46.90	69	32.23	2.41	28.00	39.00
hct_{14} (%)	44	33.28	3.46	26.40	41.20	66	31.70	2.47	26.00	37.00
hct_{30} (%)	44	29.23	3.60	22.60	38.80	58	27.36	3.86	17.00	34.00

8.3.2 QTL detection and allelic effect

Updated results of the QTL analysis using new SNP positions and the updated QTLMAP software are listed in table 8.2. Two models were considered for the QTL analysis, either considering polymorphisms appeared before breeds divergence (without breed clustering of haplotypes) or after breeds formation (clustering of the population haplotypes according to their breed origin, MBB or RMN). Modifications resulted in minor changes in QTL detection analysis.

Table 8.2: Results of the QTL detection analysis in the BC population

	GWAS^a		**GWAS**^a_b	
Trait	Pos^b	Flanking snps	Pos^b	Fanking snps
dpeps1	36.73	s62516 - s28563	38.73	s28563 - s26955
dpeps2	36.73	s62516 - s28563	36.73	s62516 - s28563

a:GWAS, association analysis without considering breed origin of the haplotype; GWASb, association analysis with breed clustering of haplotypes. *b*: Position in Mbp

Both models resulted in the same region around 37 Mbp maximizing the likelihood that a QTL affects pepsinogen variation under infection (see table 8.2). $GWAS_b$ model also pointed a QTL signal in the same region (38.73 Mbp) for dpeps1 only.

Three out of the four SNPs composing the haplotype considered for analysis, were common to both maximal likelihood positions (s62516, s28563 and s26955) suggesting a unique QTL (see table 8.3). Interestingly, only one allele was identified in both the MBB and RMN population (AAAG allele at 36.7 Mbp, see table 8.3). Hence, these MBB-specific alleles were not considered

Table 8.3: Allelic effect of the QTL affecting pepsinogen concentration estimated in the BC population

			dpeps1		dpeps2	
Analysis[a]	**Haplotype**[b]	**Frequency**	**Effect**[c]	**S.E.**	**Effect**	**S.E.**
	A**AAA**	0.08	-0.53	3.18	-2.01	2.85
	A**AAG**	0.08	-0.01	3.18	-1.31	2.76
	A**GAA**	0.05	-0.42	3.78	-1.52	3.38
GWAS	A**GGA**	0.23	-0.21	3.09	-1.79	2.72
(36.73 Mbp)	A**GGG**	0.15	-0.24	3.01	-1.51	2.64
	G**AAA**	0.1	-0.40	3.13	-1.49	2.73
	G**GAA**	0.13	-1.09	3.08	-2.55	2.70
	G**GGA**	0.09	-0.04	3.31	-1.92	2.85
	AAAA_{RMN},	0.05	-0.41	1.01	-	-
	AAAC_{RMN},	0.12	-0.16	0.20	-	-
	AAGA_{MBB},	0.25	1.01	1.09	-	-
GWASb	**AGA**A_{MBB},	0.5	-0.18	0.13	-	-
(38.73 Mbp)	**GAA**C_{RMN},	0.17	-0.96	0.21	-	-
	GAGA_{MBB},	0.25	0.07	0.49	-	-
	GGAA_{RMN},	0.05	0.35	0.71	-	-
	GGAC_{RMN},	0.27	0.07	0.18	-	-
	GGGA_{RMN},	0.11	-0.03	0.13	-	-
	A**AAA**$_{RMN}$,	0.08	-	-	-2.09	2.63
	A**AAG**$_{MBB}$,	0.25	-	-	-0.26	3.24
	A**AAG**$_{RMN}$,	0.07	-	-	-1.60	2.57
	A**AGA**$_{MBB}$,	0.25	-	-	-1.78	2.63
	A**GAA**$_{RMN}$,	0.05	-	-	-1.56	3.17
GWASb	A**GAG**$_{MBB}$,	0.25	-	-	-1.46	3.11
(36.73 Mbp)	A**GGA**$_{RMN}$,	0.23	-	-	-1.90	2.52
	A**GGG**$_{RMN}$,	0.15	-	-	-1.63	2.49
	G**AAA**$_{RMN}$,	0.1	-	-	-1.56	2.53
	G**AGA**$_{MBB}$,	0.25	-	-	-1.89	2.86
	G**GAA**$_{RMN}$,	0.13	-	-	-2.56	2.54
	G**GGA**$_{RMN}$,	0.09	-	-	-1.33	2.84

a: GWASb, association analysis considering breed origin of haplotype; GWAS, classical association analysis; associated LRT_{max} position are provided in brackets. *b*: SNPs common to each haplotype are indicated in bold ;*c*: in phenotypic standard deviation

in the GWAS analysis as their representation fell down below the 5% threshold.

This is noteworthy, provided the haplotype associated to the highest pepsinogen concentration increase came from the MBB breed (see table 8.3).

8.3.3 QTL effect on pepsinogen concentration in the BCxBC flock

To further investigate this QTL region, allelic groups were built within the whole BCxBC flock based on the GWASb analysis results for which haplotype exerted the most extreme effects. The two alleles with most extreme effects served as tag for constituting genotypic classes (table **??**). The HR group was composed of animals carrying the $AAGA_{MBB}$ allele, while LR animals carried the $GAAC_{RMN}$ allele (table 8.4). As a limited number of sheep carried the $AAGA_{MBB}$ allele (n=8), individuals carrying the $GGAA_{RMN}$ were added to this group (table 8.4). Animals carrying one allele of each defined class, or none of these were discarded.

Table 8.4: Repartition of allelic carriers in each BCxBC

QTL21	**Flock**[a]	**Frequency**
HR[b]	1	9
HR	2	28
LR	1	23
LR	2	18

a: 1, animals transferred at the INRA Langlade experimental unit; 2, animals remained at the La Sapinière experimental unit. *b*: HR, High responder, LR: low responder

Basal pepsinogen values measured before experimental infection were low and similar in both allelic groups (p=0.16, see table 8.5). Pepsinogen concentration increased until 14 dpi before getting back closer to basal level at the end of the experiment (table 8.5). This variation at 14 dpi was significantly higher in the HR group (p=0.002). $Peps_{30}$ almost fitted the basal pattern of pepsinogen concentrations (table 8.5) without any significant differences between groups for dpeps2 (p=0.13).

Negative phenotypic correlations were observed between peps1 and FEC21 (-0.34, p=0.004) and between peps1 and FEC30 (-0.24, p=0.05). However, no significant differences were found between FEC21 and FEC30 between the HR and LR groups (table 8.5).

Surprisingly, the HR group demonstrated higher resilience as assessed by the higher average hct_{14} and hct_{27} (p=0.002 and 0.01 respectively). This finding was also completed by a comparison of WB performed within flock 1 subset (n=9 and 23 for the HR and LR groups respectively).

Table 8.5: Comparison of allelic groups on pepsinogen, FEC and hematocrit

		HR			LR		
Variable	**N**	**Mean**	**Std**	**N**	**Mean**	**Std**	**p-value**
$peps_0$ (mU/L)	24	0.47	0.13	27	0.38	0.16	0.16
$dpeps_1$ (mU/L)	24	1.38	0.55	27	0.87	0.42	$1.90x10^{-3}$
$dpeps_2$ (mU/L)	24	0.74	0.28	27	0.59	0.31	0.69
FEC21 (eggs/g)	36	2,526	3,157	34	1,560	2,579	0.29
FEC30 (eggs/g)	36	12,394	8,672	34	12,687	13,090	0.23
hct_0 (%)	37	33.09	2.83	41	34.09	5.45	0.32
hct_{14} (%)	37	31.33	2.43	38	33.70	2.66	$1.70x10^{-3}$
hct_{27} (%)	36	27.03	3.68	34	29.95	3.24	0.01

$peps_0$: basal pepsinogen concentration; $dpeps_1$: variation of pepsinogen concentration at 15 dpi; $dpeps_2$: variation of pepsinogen concentration at 30 dpi; FECX: FEC at X dpi; hct_X: hematocrit at X dpi

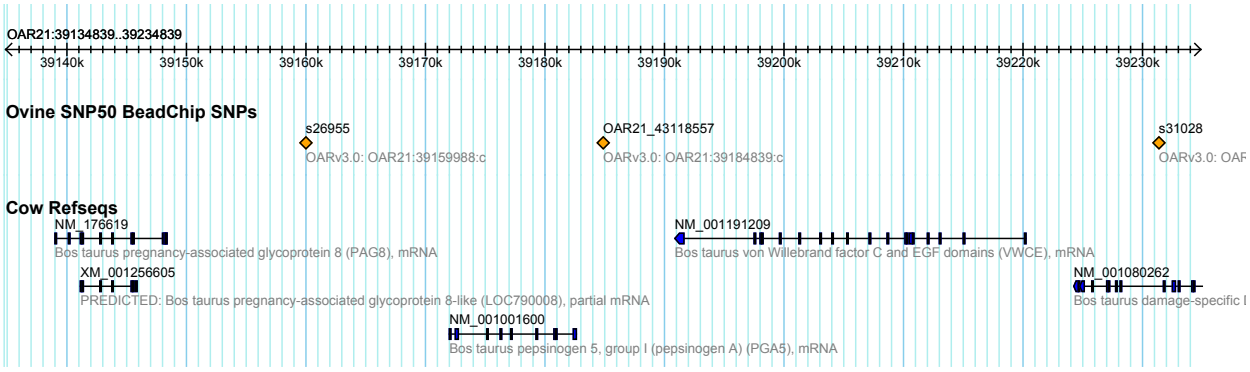
This test showed that HR animals also had significantly less worms established (p=0.03).

The two allelic groups showed similar patterns for other traits (table 8.5).

8.3.4 Sequencing the *PGA5* locus

Applying a GWASb model to the BC data identified a 4-SNP haplotype associated with pepsinogen concentration variation at first infection. Interestingly, the two last SNPs of the haplotype *i.e.* s26955 and OAR21_43118557, flanked the *PGA5* locus (see figure 8.1). Given that this gene encodes the pepsinogen protein, sequencing was performed in two animals belonging to each of the two considered allelic groups. Preliminary results of this investigation are reported.

Figure 8.1: View of the ovine sheep genome between the s26955 and OAR21_43118557 SNPs



The *PGA5* was sequenced in two animals of each allelic group. No SNP could be found in exons. On the contrary, 32 SNPs and 2 insertion/deletion events were unraveled in intronic sequences, nine of which being differential between the two allelic groups (table 8.6).

Table 8.6: No. markers found in intronic sequence of *PGA5* between the allelic groups

	SNP	**In/Del**[a]
Shared	22	1
Group specific	8	1
Total	32	2

a: Insertion/Deletion

8.4 Discussion

This study aimed at validating a QTL affecting serum pepsinogen during GIN infection. This was achieved by comparing allelic carriers of this QTL and by looking for particular marker within a candidate gene.

So far, pepsinogen concentration has been used in veterinary medicine as an indicator of flock infection burden by worms living in abomasum, *i.e. H. contortus* and *T. circumcincta* in sheep [340, 507, 164, 472] and *O. ostertagi* in cattle [380]. The higher serum pepsinogen, the higher worm burden. Still, pathophysiological mechanisms explaining this rise are not clear. Especially, it is difficult to determine if this rise is mediated by the parasite itself or occurs as collateral damage because of the host response [471]. In addition, Terefe *et al.* found that the resistant MBB sheep breed had higher serum pepsinogen than the susceptible RMN breed throughout infection, hence suggesting a putative favorable relationship between pepsinogen and infection outcome [507]. The previously estimated correlations between pepsinogen concentration and FEC at reinfection agreed with this hypothesis [454]. This result was also found in the BCxBC crosses during their primary challenge. This suggests that abomasal tissue damage may be concomitant of a reduction in eggs excretion, and putatively to decrease worm burden.

Results from the previous linkage analysis also linked FEC and serum pepsinogen together with two QTL overlapping the 35-40 Mbp region [454]. This QTL was subsequently confirmed by the association analysis for pepsinogen concentration but not for FEC. In that case, one of the 4-SNPhaplotype segregating in the RMN breed was associated to the lowest rise in serum pepsinogen after infection, whereas the opposite extreme allele was found in the RMN breed. This finding is in good agreement with results reported by Terefe *et al.* who measured a 3-fold difference in serum pepsinogen between the MBB and the RMN breed [507]. Still, our findings also suggest that some RMN allele could contribute to a significant rise in pepsinogen concentration under infection. This was confirmed by the comparison of HR and LR groups that

revealed a 0.82 phenotypic standard deviation. This is somewhat lower than the 1.4 deviation predicted by the QTL detection estimation, but still indicates a significant difference that reinforce previous results. In this case, no correction for the effect of the rest of the genome was performed. Indeed, two few pepsinogen data had been recorded to correctly estimate gEBV for this trait. A correction could have been done by considering every other QTL affecting one of the pepsinogen-related traits (n=5). However the related increase in statistical levels to be tested did not match the sample size we had. A putative effect of the rest of the genome would have impacted differences between groups but should not modify the observed trends, as selection of BCxBC animals was random at every region other than OAR12.

Interestingly, the HR and LR groups exhibited same basal serum pepsinogen (p=0.16), hence suggesting the QTL allele they carried affects pepsinogen concentration only after GIN challenge. The *PGA5* locus was of primary interest as its function directly matches the trait associated to the QTL. Therefore, sequencing was attempted that revealed several markers segregating within intronic sequences. Some of these SNPs were differential between the couples of animals belonging to the HR and LR groups. Still these encouraging findings needs to be confirmed in more animals of each group. Additional functional work will also be required to determine any alternative splicing occurring and/or post-transcriptional modifications.

Lower blood losses were observed in the HR group and less worms were retrieved from the subset of necropsied HR lambs. Hence, the QTL region may not only control serum pepsinogen during infection, but it could also affect the outcome of *H. contortus* infection, as already outlined by the linkage analysis performed in the BC population. This broader function of the QTL region had also been suggested by Dominik *et al.*'s (2010) who found a QTL associated to eosinophils counts that overlapped this region [128].

Differences observed for hematocrit and WB between HR and LR animals may reflect the action of the rest of the genome. HR animals could be more resistant on average than their LR counterparts. However, selection was done at random on the rest of the genome and the two groups showed similar gEBV for resistance to GIN (p=0.26, gEBV estimated using FEC at first infection and the whole BC and BCxBC population).

A maybe more realistic explanation could lie in the chromosomic region we are interested in. Indeed, the targeted QTL lies within a gene-rich segment that encompasses among others, the CD5 and CD6 genes (located at 36.6 and 36.4 Mbp respectively). These genes are known to act on T-cells differentiation and regulation [116, 398]. Assuming 1 cM is equivalent to 1

Mbp, approximately 3% recombination are expected to break down the genomic relationship between the *CD5/CD6* and the *PGA5* locations. Therefore, transmission most often occurs without recombination. Hence, it might be possible that the observed 4-SNP allele associated to serum pepsinogen rise is in disequilibrium with favorable alleles of the *CD5/CD6* genes. Favorable *CD5/CD6* genes might contribute to a better Th2 response, hence resulting in HR animals being more resistant and producing more pepsinogen during infection. These two latter phenomena could be either independent or cause and consequence. Interactions between the two regions are also conceivable and sequencing the CD5 and CD6 genes in HR/LR lambs should help confirming this hypothesis.

If confirmed, this hypothesis should have practical consequences on the interpretation of pepsinogen concentration during GIN infection, as pepsinogen could tag individuals exerting the most efficient immune response.

8.5 Conclusion

Our study was designed to validate differences in serum pepsinogen predicted by a QTL genotype mapped on OAR21.

Two groups of BCxBC, *i.e.* high- and low-responders, were built based on their QTL allele. Our results showed that this QTL region predicted well pepsinogen concentration rise after infection. It also suggested that this QTL affects circulating pepsinogen in an inducible fashion as no difference in basal circulating pepsinogen could be observed. Sequencing of the *PGA5* locus in two couples of animals belonging to each of the allelic group revealed 9 specific markers in intronic sequences. Additional genotyping of remaining animals should confirm/infirm these markers.

Our findings also suggested the QTL under investigation could affect the outcome of GIN infection. No obvious explanation can be proposed. Still, the close relationship between the CD5/CD6 loci and the PGA5 locus might explain colinearity between the QTL allele effect on serum pepsinogen and infection outcome.

This latter hypothesis requires further investigations before being confirmed. In case, PGA5 markers were confirmed, further study would be required to check for any alternative splicing and/or post-transcriptional event.

Chapter 9

Mixing previous microarray results and genome scan findings: genes network analysis

9.0.1 Rationale

In the frame of a previous study [316], a microarray expression analysis has been performed in pure breed animals to investigate any differential expression in functional candidates during early response to *H. contortus*, *i.e.* 8 days after challenge. From this study, a few genes were found to be differentially expressed between MBB and RMN infected individuals [316]. Thanks to the annotation effort of the ovine genome, candidates can be positioned along the genome thus allowing for comparison with the genome scan hits obtained.

In addition, a gene network analysis merging both functional candidates differentially expressed in pure breed animals and the annotated genes within the QTL confidence intervals should provide additional insights that could help identifying genes involved in resistance to *H. contortus*.

9.0.2 Reminder: microarray study in pure breeds (Liénard *et al.*, manuscript in preparation)

A microarray analysis has been performed to compare gene expression between eight MBB and eight RMN individuals challenged with 12,000 L3 larvae of the IseFb *H. contortus* strain. Four additional animals of each breed served as uninfected control. Eight days after infection lambs

were sacrificied by intra-venous lethal injection and abomasal lymph nodes and abomasal fundic mucosa were sampled for subsequent RNA extraction (kit RNeasy purification, Qiagen). After RNA quality control, RNA were tagged with fluorescent dyes before being hybridized on the Agilent ovine microarray.

Microarray data have been normalized and analyzed with the BioPlot software (http://biopuce.insa-toulouse.fr/ExperimentExplorer/doc/). Among genes with significant fold change (above 1.5), 157 genes were differentially expressed in one of the two considered tissues between infected groups and had HUGO reference. These genes were considered for subsequent analysis.

9.0.3 Preparation of the gene list based on QTL regions

The second list of genes was prepared based on the QTL analyses results. To minimize the risk of considering false positive signals, any QTL that had reached the 1% genome-wide significance threshold in the LDLA analysis was considered (n=38). Using the annotated list of genes available on the sheep genome browser (http://www.livestockgenomics.csiro.au/cgi-bin/gbrowse/oarv3.0/) and considering the highest range of LD in sheep, *i.e.* 4 Mbp, every annotated gene lying within +/- 4 Mbp around the position of maximal LRT value was added into the list. In the end, 1410 annotated genes with referenced HUGO names were considered.

9.0.4 Analyses performed

The two lists of functional and positional candidates were merged together and fitted into a gene network analysis using the WEB-based GEne SeT AnaLysis Toolkit (WEB-GESTALT, http://bioinfo.vanderbilt.edu/webgestalt/) [577]. Three different analyses were performed, namely gene ontology, KEGG and PATHWAY analyses. In order to minimize false positive, networks containing at least 10 genes (maximal setting permitted) were considered. For the pathway commons analysis, the most restrictive adjusted p-value ($p < 10^{-6}$) was considered after applying the Benjamini-Hochberg correction. A more conservative adjusted p-value of 10^{-4} was fitted for KEGG analysis, allowing other functions than metabolic pathways to be detected. In the same way, a gene ontology regrouping at least five genes with an adjusted p-value below 5% was considered significant.

9.0.5 Results

9.0.5.1 Intersecting candidates

Merging the two lists of functional and positional candidates resulted in nine genes being found in common in both analysis (see table 9.1). Interestingly, the *ENO-1* gene located at 42.4 Mbp on OAR12, *i.e.* within the confidence interval of the QTL region under study, was also over-expressed in the MBB breed at 8 dpi. Interestingly, this gene lies 153 kbp away from the first SNP of the region under selection in the MBB population.

Table 9.1: Intersecting candidate genes

HUGO name	OAR	Upstream position (bp)	Downstream position (bp)
SLC31A1	2	10,576,623	10,611,739
PTGR1	2	11,995,427	12,024,932
ENO1	12	42,405,213	42,416,823
DBNDD2	13	73,124,142	73,126,403
ACD	14	34,360,043	34,366,417
SWAP70	15	42,746,505	42,822,969
HTATIP2	21	23,998,425	24,011,606
ASRGL1	21	38,501,656	38,522,679
FRAT2	22	17,890,660	17,891,299

9.0.5.2 Gene ontology

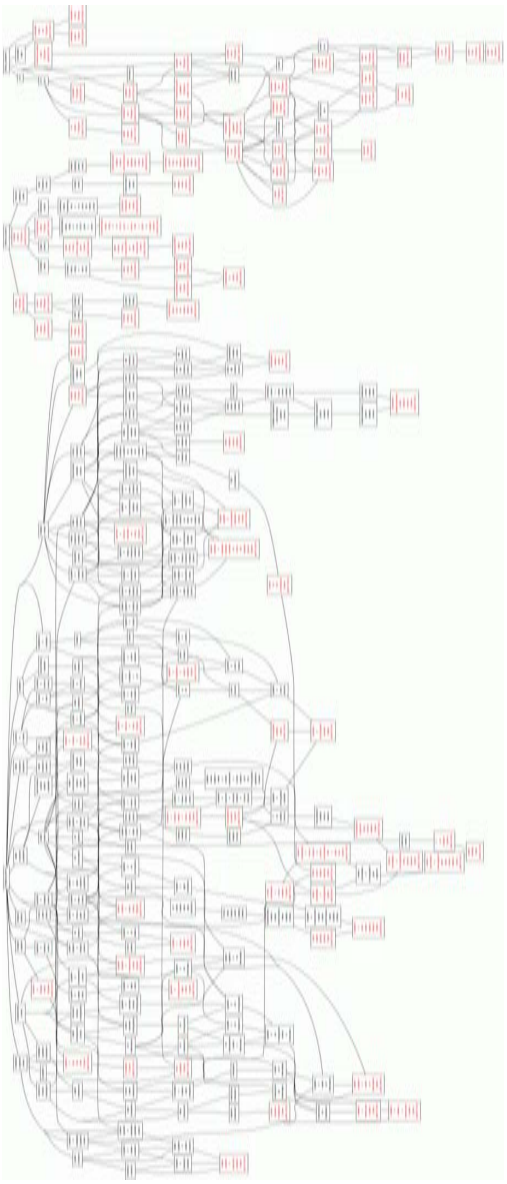
A relatively loose threshold has been chosen for gene ontology analysis, *i.e.* at least five genes with the same ontology with an adjusted p-value of 5%. Any analysis with more restrictive parameter resulted in few ontology terms mostly related to molecular functions, *i.e.* catalytic activity (457 genes, p=7.10^{-3}) and catabolic process (698 genes, $1.4.10^{-4}$), followed by the SMAD-protein complex assembly (7 genes, p=7.10^{-4}) and lysosome formation (30 genes, p=5.10^{-3}) for the biological process and cellular components classes (data not shown).

At the 5% significance level, many more ontology terms were found (figure 9.1), among which the immune function, phospholipase activity and blood coagulation terms could be directly related to haemonchosis (see figure 9.1).

9.0.5.3 KEGG and metabolic pathways analyses

Results of both analysis were put together. The immune response associated functions was among the top three pathways and represented 14% (see figure 9.2), being almost as high as

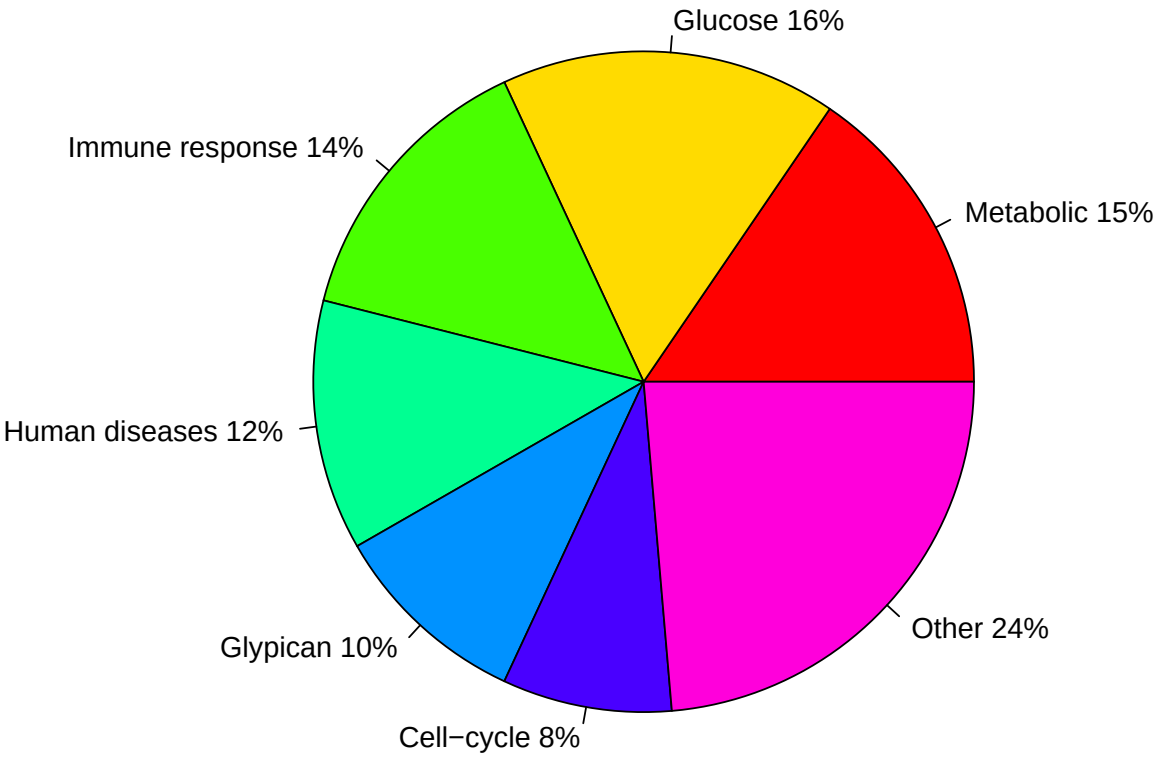
Figure 9.1: Results from the gene ontology analysis[a]



a: This figure is viewable in the generated .pdf file after applying a 800% zoom in

glucose related (16%) and metabolic pathways (15%).

Figure 9.2: Relative proportions of the metabolic pathways



9.0.5.4 Over-representation of some chromosomes

After analysis, some chromosomes were more frequently encountered in the different metabolic pathways. For instance, 174 and 100 genes located on OAR14 and OAR21 respectively, were involved in one of the considered pathways hence reprenting one third of the total networking. Chromosomes 22, 12, 23, 17, 15 each represented 5% of the total number of genes involved while remaining chromosomes carried less than 1% of the total genes.

Given that resistance to GIN should involved genes related to the immune function, frequency

Table 9.2: Chromosomes representation in immune related functions (Pathways and gene ontology analyses)

OAR	Frequency in pathway analysis	Frequency in Gene Ontology
2	2	-
4	1	-
5	2	-
7	3	4
8	1	1
9	1	1
12	8	9
13	3	3
14	13	6
15	1	5
16	1	-
17	5	2
21	9	4
22	2	3
23	1	3
26	1	2

of occurence of chromosomes was computed by considering immune-related function (Chemokine signaling pathway, TNF receptor signaling pathway, BCR signaling pathway, Bcell receptor, Tcell receptor, TCR signaling in CD4+Tcells, signaling in immune system). Considering these pathways, OAR14 ranked first again representing 25% of the occurences (table 9.2). OAR12 and OAR21 were the two followers (8 and 9 occurences respectively).

Similarly, genes related to immune-related GO terms (*.e.g* immunoglobulin production during immune response ...) and phospholipase activity and blood coagulation were most frequently encountered on OAR12 (n=9) and OAR14 (n=6) (see table 9.2).

9.0.6 Discussion

The purpose of this investigation was to identify candidate genes intersecting a micro-array experiment performed in pure breeds and the QTL detection study, and to target candidate genes explaining detected QTL. In this latter case, such analysis, even if highly dependent on previous existing knowledge, provides a good way to establish relationships between detected QTL while avoiding computation issues of including epistasis effects in QTL detection model.

Such work has been recently published by Sayre *et al.* (2011) who performed a gene network analysis mixing QTL knowledge (QTL related to parasite resistance in mice, rat, sheep, cattle

and human) and a gene set from microarray experiment performed in sheep [460]. Analyzing both datasets together identified 14 pathways among which four were related to *MHC* and *IGN*γ genes. These findings helped targeting candidate genes common to every tested parasite-host species. However interpreting these results is particularly hazardous given the wide range of diversity among the considered parasites. Indeed, the authors not only considered several nematodes from different clades but also trypanosomes and leishmanies whose interactions with their hosts are completely different [460]. Our study differs in that, same MBB/RMN genetic background was considered in the micro-array analysis (pure breeds) and the QTL detection study (back-cross). In addition, the analyses were specific of the same *H. contortus* strain.

This study has been performed using the human genome as reference set. But annotation of the ovine genome is under completion. If this annotation is biased towards particular gene functions, this might hence induce a bias in the subsequent analysis. For instance, if more genes are related to immune function, then QTL regions could be more often associated to immune pathways, thus missing true associated genes. Another analysis should be performed with the complete gene list of the ovine genome to take this parameter into account. For time purpose, this has not been done at the time of writing this manuscript.

Still, a suprisingly limited number of genes were found in common between the differentially expressed genes from micro-array and the detected QTL. It is worth that among these genes, ENO1 was located very close to the sweep signal detected in the MBB breed. In case a functional role in resistance to GIN could be associated to the sweep, this gene constitutes a strong candidate. Indeed it encodes alpha-enolase that has been demonstrated to reduce the release of TNF_{α} by mast cells in mice [447]. It has also been involved in strong inflammatory response in rhumatoid arthritis when expressed on macrophages and monocytes surface [27]. The lack of intersecting candidates between the two studies may be due to the different time of infection considered. The micro-array experiment focused on early steps of the GIN reaction by measuring gene expression at 8 dpi. On the contrary, BC sheep were sampled at 25 and 35 dpi and precise phenotyping had been conducted 45 days after the second infection.

Interestingly, three chromosomes were particularly involved in immune function. OAR14 was highlighted as a central player whereas it carried a weak QTL signal in the BC population. OAR12 and OAR21 whose properties are particularly investigated in this PhD work were also often involved in immune pathways and phospholipase activity which is directly involved in inflammation process.

9.0.7 Conclusion

The gene network analysis allowed bringing together different research findings performed on the same genetic back-ground, *i.e.* MBB and RMN breeds.

A limited number of genes was found in common between experiments certainly due to the different time of infection considered.

OAR12 and OAR21 appeared to be two important players in response to *H. contortus*. In addition, OAR14 was preponderant in identified pathways whereas it seemed to carry only a minor QTL.

Part V

Discussion and perspectives

9.1 A brief overview of the results

QTL mapping analyses highlighted several regions whose variability could be worth further studies and eventually selection. Among these, two QTL regions were thought of outstanding interest and deserved additional investigations during this PhD project.

On one hand, OAR12 carried a 10 Mbp region that was common to two QTL affecting FEC at first and second infection respectively. The QTL affecting FEC in naïve lambs exhibited a $0.19\sigma_p$ effect. An additional finding from a population genetics approach demonstrated that a 5-SNP region was fixed in the MBB pure breed but still segregated within the RMN population. A validation study was implemented by selecting BCxBC animals based on their sweep genotype. This region did not explain much of the observed variation between BCxBC groups. The additional phenotyping performed in these BCxBC animals highlighted the QTL region effect on female worm fertility. Further, concomittant stronger Th-2 cytokinic environment was found in carriers of one MBB allele associated to resistance as opposed to the IBS-RMN allele carriers. Questioning thus arises to determine if this region under selection matters or not, and whether the physical proximity could prevent the delineating of this question. In addition, no obvious candidate gene has been identified due to the width of the region.

On the other hand, OAR21 was significantly associated to pepsinogen concentration which has been used in veterinary medicine to monitor infection burden at the flock scale. The position maximizing the likelihood that a QTL exists mapped the *PGA5* locus which codes for the pepsinogen. Validation of this functional candidate was undertaken by sequencing exons and introns in a subset of BCxBC individuals identified as either "high" or "low" pepsinogen responders. Results confirmed expected differences in pepsinogen variation under infection and nine markers were identified as specific of each designed group. Not only showing differences in pepsinogen concentration, allelic groups exhibited significant differences in worm burden, and favorable correlations were found between pepsinogen variation and FEC in both BC and BCxBC populations. It is difficult to draw any conclusion on this point, but the presence of *CD5* and *CD6* loci in the vicinity of *PGA5* deserve further attention. Indeed, both genes are known to act as T-cell regulators [398] and their close location to *PGA5* is in favor of a co-inheritance or inter-regulations or both.

Interestingly, a gene network analysis based on the QTL knowledge showed that OAR12 and OAR21 were particularly involved in immune-related pathways. It also put OAR14 as a central

region also being involved in immune functions. A weak QTL signal had been found in the BC population on this chromosome, but findings of other research teams (significant QTL affecting FEC in a Scottish Black Face population, S.C. Bishop, personal communication and a selection sweep identified in a New-Zealand sheep population, M.I Fariello, personal communication) also tend to rank this chromosome among top players in resistance to GIN infection.

9.2 BC design, an old-fashioned tool ?

9.2.1 Purpose of the BC design

An abundant literature has been produced on the contrasting pattern of susceptibility to GIN infection between the MBB and RMN breeds [198, 24, 507]. Crossing the two breeds merged the two respective genomes into F1 sires that were subsequently back-crossed to RMN ewes and finally resulted in creating huge families of approximately 250 progenies. Knowing the pedigree structure, it was hence possible to compare the effect of the MBB genome to the effect of the RMN genome at every marker location. Thus, the more contrasted the phenotypes between original breeds, the more efficient is the QTL detection analysis. Crossing these two breeds was thus of primary interest for investigating the genetic background explaining the observed variation in the resistance to GIN infection. For implementing this design, a limited number of MBB sires were available at the time of experiment so that only five MBB sires were chosen at the beginning of the QTL detection process, and four MBB haploids finally remained in the 50-K dataset. This practical issue was synonymous of a poor screening of the genetic variability within the MBB breed. This point was not thought of as an issue provided the working hypothesis was to consider the MBB breed as resistant, the putative QTL being fixed in this breed. Indeed, none of the studies had pointed out a MBB sheep being as susceptible as the average RMN sheep under haemonchosis. The MBB breed thus served as a basal pattern of resistance on which to oppose the RMN alleles whose variability was higher. This hence resulted in screening the RMN genetic diversity for particularly susceptible alleles.

In the end, this strategy was successful as numerous significant QTL have been identified among which a limited number of regions exerted higher effects on OAR5, 7, 12, 13 and 21. The MBB alleles exerted the most favorable effects on measured phenotypes. The 50K SNP chip provided new information that permitted the reduction of QTL confidence intervals and unlocked the QTL scanning on some neglected chromosomes like OAR21.

However, one reviewer of the QTL mapping paper [454] stated : " *The experimental population was designed for an "old-fashioned" linkage study and, therefore, it is questionable how appropriate the design is for SNP (LD and LDLA) analyses.*". Indeed, BC population was created almost ten years before the recent progresses of genomics and the release of the ovine SNP chip. Practical implications are discussed below.

9.2.2 The back-cross design helped exploring the RMN breed diversity

The comment proposed by the reviewer does not hold for the RMN population. Indeed, more than 600 RMN ewes have been used for creation of the BC flock. Even if the only F1 chromosomes could be investigated in the within-analysis, LD-based analysis focused on every available chromosomes within founders. Implementing a breed clustering of the segregating alleles hence gave access to the effect estimation of every RMN allele.

This increase in the number of degrees of freedom usually resulted in increasing the overall likelihood (see [454]). This was particularly true for the QTL affecting FEC in naive lambs on OAR5, and FEC at reinfection on OAR13. This is in favor of a great within-breed variability, that could be explored with the back-cross design. Implementing such a F1xRMN back-cross was hence useful as it screened the RMN population for particularly unfavorable alleles segregating in the high-producing breed. Identification of these alleles and confirmation of their effect within the RMN breed is particularly interesting, as their elimination in the frame of a breeding scheme should be easier to implement than the introgression of any particularly favorable MBB allele into the RMN breed.

On OAR12, the inclusion of the RMN maternal alleles in the analysis resulted in a second peak appearing around 10 Mbp. According to the analysis implemented, this peak could also reach the highest likelihood, hence shifting the QTL position or at least increasing the QTL confidence interval and making the study of the QTL region more difficult. This illustrates the different level of genetic variability that were investigated, that focused either on the within breed variability or on the comparison of two breeds genetics.

9.2.3 Practical consequences of the limited number of MBB sires

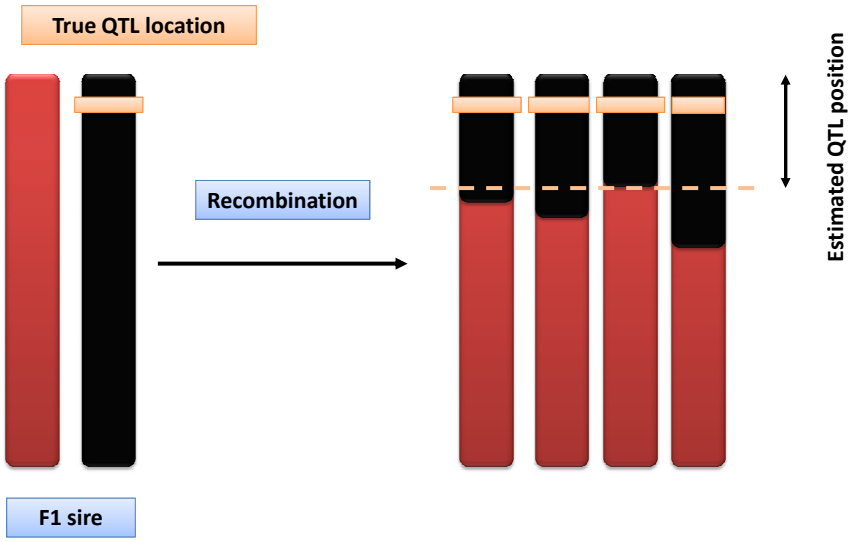
The back-cross design permitted a good screening of the RMN population alleles. However, only four out of the five original F1 sires were genotyped with the 50K SNPchip.Thus, the contribution of MBB alleles to resistance to GIN was poorly explored. In addition, the sampling of four MBB

sires may have biased our analysis towards common variants of the MBB breed. Provided the MBB breed is more resistant than the RMN breed on average, resistance determinism may lie within these common variants and focusing on these should not prevent the identification of QTL regions explaining between breeds differences.

However, this analysis might not reflect the probable variability existing in the MBB breed and thus provides a limited access to its ancestral LD. Indeed, performing an association analysis within the MBB population in the BC design was equivalent to a statistical comparison of the four chromosomes and their associated phenotypes (that are the average phenotype of the progenies that inherited these chromosomes). This was not expected to bring much more additional power to the QTL detection analysis.

Therefore, the exploitation of the MBB genetic material was mostly achieved through the LD generated by the back-crossing process. This recent LD spans long interval which is thus expected to limit the precision of the QTL mapping analysis. Indeed, during back-crossing process, relatively large fragments of chromosome recombine. Let's consider a QTL, fixed in one breed and located at the top telomeric end of the chromosome (see figure 9.3). Recent recombinations occurring during back-crossing generate recombinant chromosomes (red and black, figure 9.3) that differ in the size of the recombinant fragment they carry (black fragments on figure 9.3).

Figure 9.3: Back-crossing and precision of QTL mapping



Assuming genetic markers are available on this black fragment and are in sufficient LD with the QTL, estimated likelihood that a QTL is present will be roughly equivalent all over the smallest black fragment common to every recombinant chromosome. Provided only one recombination event has occurred, this estimated position will span relatively wide interval, no matter how many markers are available. Therefore, using a back-cross population for QTL detection purpose is not particularly well-suited for taking advantage of SNP markers density.

In that way, SNPs data did not bring much more useful information about the MBB breed, making the BC design an "old-fashioned" tool.

9.2.4 Back-crossing is synonymous of multiple sources of LD

Back-crossing is synonymous of mixing two populations with diverse genetic background. This admixture effect has been particularly investigated in human genome-wide association study, as not taking people's ethnicity into account results in false positive association (already discussed in introduction 5.3.1).

In this study, assuming that every marker comes from the same breed results in tagging mutation appeared before breeds diverged. In that case, it is questionable whether this mutation can still be detected. Indeed, if it confers an evolutionary advantage it might come to fixation after a limited period of time. On the contrary, if this mutation is unfavorable to fitness, it might have been eliminated from the population. From a technical matter, such ancestral mutation would have been passed on through many generations hence multiplying the number of recombinations and thus reducing the associated LD. Therefore the SNP density achieved on the 50K SNP chip would not be high enough to pinpoint these mutations. Using this 50K SNP chip hence questions more recent mutations and not considering a breed origin of the haplotype is equivalent to consider that the marker and the QTL linkage phase have followed an equivalent evolution. This was clearly not the case, as illustrated by the estimation of the non-persistence of the LD phase across both breeds [454]. This point has also been pointed out in multi-breed genomic prediction attempts [103].

One way to correct for this population stratification was to modify the association model implemented in the QTLMAP software, in order to differentiate IBS haplotypes based on their breed origin (C. Moreno, personal communication).The resulting likelihood was a composite of an association performed in the MBB breed and another one led in the RMN breed (C. Moreno, personal communication). Implementation of this model also contributed to confirm that LD

phase did not hold across breeds as some IBS QTL alleles did not exert same effects.

Given the two populations do not exhibit same LD pattern, *i.e.* LD decreasing faster according to physical distance in the RMN breed [454], it is questionable whether the MBB-associated likelihood profile cannot artificially increase the resulting QTL confidence interval. Indeed, the MBB breed shows higher level of LD than the RMN breed. Therefore, some marker-QTL associations should arise at a larger distance in the MBB breed than in the RMN breed. This has been illustrated by considering the LD pattern on OAR12 within pure breeds and the BC populations (see 6.2.3 and figure 6.1). Combining likelihoods of the association performed in the MBB and the RMN breeds might thus combine a peaky RMN-breed likelihood profile due to the rapid decline of LD in this breed to a wider likelihood profile provided by the MBB association. This point remains difficult to investigate within our design but could be studied by simulations. A QTL could be simulated to segregate within each breed before running a QTL analysis on each pure breed subset. Following this pure breed analysis, a BC design could be simulated as well based on the QTL properties simulated within each breed, in order to properly model its expected frequency, LD between markers and the QTL according to the breed origin and its allellic effects. Subsequently, a second QTL analysis should be performed on this dataset and compared on the results obtained within each breed to assess how the back-crossing impacts the likelihood profile.

Considering the full impact of admixture in this design is complicated by the heterogeneity underlying each of the two considered breeds : the MBB breed has been composed of various local breeds from West-Indies while the RMN breed has been created from the Berrichon du Cher and the Romanov breeds in late 70's. Such genetic melting should contribute to the unraveling of QTL for GIN resistance by benefiting the within-breed genetic diversity, but it might complicate any marker-assisted selection as the linkage phase between markers and QTL may not hold within breed.

9.2.5 What could have been done ?

9.2.5.1 Performing a F1 x MBB back-cross

The BC population has been created in two successive steps. A first BC has been created and extreme BC individuals have been measured for additional patho-physiological phenotypes measured on dead animals. Hence, the most interesting animals were not available any more for further validation study. Therefore a second campaign of matings was implemented to confirm

the QTL signal in another pool of BC offsprings and to produce sufficient BC individuals to perform subsequent marker-assisted matings for in-depth functional investigation. In addition, the creation of this second BC population also allowed to screen for QTL affecting pepsinogen concentration.

A good complementary experiment to this BC design might have been to redo a back-cross of the F1 sires on MBB ewes. Increasing the available number of MBB haplotypes would have benefited the LD-based analysis by exploiting ancestral LD. by better exploring the MBB breed diversity, hence unraveling putative QTL at low frequency.

9.2.5.2 In case of unlimited research funds ...

The BC flock was created before the SNP chip was released. Therefore, it was more pertinent to use the already available animals and their associated phenotypes, rather than setting up a new design. Even if BC design is not particularly indicated to take advantage of LD in QTL mapping, the application of SNP data with this BC design was both cheaper than a new design and successful at identifying new QTL regions and bringing more of information from the RMN maternal population.

Assuming the whole experiment should be set up today (with no limit for research credits), another framework should be implemented as proposed by Weller et al. [552]. Firstly, a QTL mapping should be performed at a population level to take advantage of ancestral LD and hence to pinpoint the QTL location. Technology now provides high-density panel of markers in sheep that should be increased to 800K by the end of 2012 (C. Moreno, personal communication). Using this type of data, it would also be possible to exploit lower level of LD and thus to target more ancient mutations. It would also enable the identification of additional selective sweeps.

Therefore, two populations with marked differences in resistance to GIN infection should be phenotyped and genotyped with a high-density SNP chip. An association analysis should be run on genotypic and phenotypic data to determine within-breed QTL affecting resistance. As the contrasting pattern observed between phenotypes of individuals belonging to the same breed might be shorter than the one observed between breeds, more animals would be required to achieve the same power as the one achieved in the BC population. Therefore more animals than the 1,000 BC lambs may be required for achieving the same detection power. These locations could be subsequently compared to target QTL segregating in both breeds. Selective sweeps could also be looked for within each breed in order to target regions differentially selected in

the susceptible or the resistant breed. Merging findings of QTL mapping analysis and of the population genetics data could help pinpointing true QTL location. The use of both datasets combined with the application of several methodologies could also contribute to reduce the number of false positive.

Once the QTL location has been pinpointed at the population level, a limited number of sires should be selected to produce progenies. Progenies would be divided into two groups, based on which paternal haplotype they inherited. Based on the differences of the two compared groups, it should be possible to determine if their sire was heterozygous or homozygous for the QTL. Trios composed of sires, susceptible and resistant progenies could be sequenced to determine the underlying Quantitative Trait Nucleotide (QTN) that would be subsequently validated in a functional validation trial based on this QTN.

9.3 About the functional investigations

9.3.1 Did we choose the right QTL ?

One major feature of this PhD project was to conduct an in-depth investigation of particularly interesting QTL region. Functional validation was performed on animals selected according to their QTL genotypes. Due to the limited time frame allowed for the project completion, selection of BCxBC parents was based on the results of the linkage analysis. After this analysis, OAR12 was the only chromosome that carried a limited region that significantly affected several traits, *i.e.* FEC in naive and immune lambs. In addition, it exhibited one of the most important estimated effect (-0.19 σ_p on FEC) and this QTL was also one of the only QTL to be detected in the second generation of BC[454].

The QTL carried by OAR21 significantly affected pepsinogen concentration. As such, it was not of primary interest as direct application for selection purposes were not obvious. Still, this region had also been associated to FEC at first infection, hence suggesting a more complex way of action than the only pepsinogen concentration variation. In addition, the association analysis rapidly offered a strong functional candidate gene whose role in the QTL signal could be easily investigated.

Other QTL could have been chosen for further investigation. For instance, OAR5 was repeatedly associated to IgG concentration and FEC in naive individuals. Still, the sole linkage analysis suggested a huge confidence interval spanning most of the chromosome hence making

it impossible to properly select animals based on this finding. Furhter, this QTL could not be reproduced in the BC2 population.

On the contrary, the QTL identified on OAR13 was of particular interest as it exhibited the highest estimated effect and the confidence interval was relatively short in the linkage analysis. These two properties would have facilitated any selection work for validation. Further, its significancy also held in the BC2 population thus reinforcing these findings. However, the only association for this narrow region was found for mean FEC at reinfection in the BC flock.

9.3.2 On the focus on first infection

In the frame of this functional validation, BCxBC animals were experimentally trickled once with *H. contortus* and differences were compared after a one-month period. The QTL affected FEC at both first and second infection and functional validation could have been performed either on naive or on immune lambs.

It seemed better to focus on the first challenge, as selecting lambs able to control parasite infection by the end of the first challenge should be protected for next infections. In addition, the fact that the investigated QTL region was involved each time BC lambs faced the parasite was also in favor of a rather innate phenomenon. The results of the micro-array experiment showing a higher induction of lectin genes and TFF3 in the MBB breed were also in favor of differences in the innate defenses, that could have matched the targeted QTL region.

In addition, Terefe et al. [507] demonstrated that major differences in immune response occurred between primed MBB and RMN lambs. At reinfection, RMN lambs were less susceptible and their pattern of resistance and cytokinic response was closer to the one of MBB lambs [507].

As a limited number of BCxBC lambs were available, splitting the flock into two groups followed up for one infection or two successive infections was not possible. Investigating two successive infections in the same lambs could have been achieved through an abomasal cannulation as it has already been reported [437]. Abomasal tissue samples could have been taken using endoscopy hence allowing a sequential analysis of both early and late time points in the same animal. However, this approach is questionable in terms of ethics and animal welfare. Further, it needs a well-skilled surgeon to avoid iatrogenic complications and the impact of the cannula on the host reaction at the site of surgery should undoubtedly interfere with the host response to parasite.

9.4 Perspectives on the study of genetic resistance to GIN

9.4.1 Should FEC be still considered as a reference trait for assessing resistance to parasite ?

FEC has been long recognized as a reference measure of the infection burden by GIN in sheep as it provides an indirect indicator of both the individual's level of resistance and its capacity to disseminate parasites [351, 130]. However this measure is an indirect indicator of the true level of resistance of the individual. In addition, sampling steps associated to FEC determination are critical in reducing precision of resistance assessment of each individual, and the applied dilution factor adds an extra-variance to the observed raw eggs count [514].

This point is critical for QTL detection study as it can hide (too few eggs counted in the whole population) or increase (lack of sensitivity for some but not all individuals) the true observed variation. This point has been discussed in detail by Bishop & Woolliams (2010) who focused on field disease data [50]. Their point was to demonstrate that genetic parameter estimation based on this kind of data could be affected by the sensitivity and specificity of the test, and by the epidemiological parameters related to the disease [50]. Their results demonstrated that suboptimal diagnoses could lead to underestimation of heritabilities and that the same imprecise phenotyping would reduce QTL detection power by a factor depending of the test specificity and sensitivity [50].

Recent improvements for FEC determination have been proposed by Mes *et al.* (2007) who proposed an automatic FEC determination by combining clear egg preparation and image analysis [360]. Authors reported higher sensivity of eggs detection, the proportion of negative animals dropping from 70% with the traditional McMaster egg isolation method to 21% with their method while the lower detection limit was increased by a 150 factor [360]. Development of such automatic methodologies with higher sensitivity should contribute to better estimate heritability and QTL effects.

Not only associated to a sub-optimal sensitivity, FEC also "summarizes" the host-pathogen interplay, as the final eggs output results of a wide range of processes linked to the nematode, to the host and to their interactions. Measuring this trait after one month is expected to differentiate between individuals with poor infection control capacities from other that withstand worm infection, hence being of interest for breeding purposes. Still, it is also equivalent to a black-box that may mask sub-categories of individuals. Thus, FEC is a good starting point to

get a rough idea of what regions are being involved in resistance to GIN without any *a priori* on finer mechanisms. In addition, it constitutes a good basis for resistance monitoring in the frame of breeding.

However, it seems necessary to refine phenotypes and to increase the number of available data for QTL analysis. This increase in available phenotypic data has been called *"phenomics"* [238]. Phenomics is thought to help understanding relationships between phenotype and genotype [238]. Last but not least, phenomics helps better defining the phenotype being investigating [238]. For instance, sheep that have developed strong hypersensitivity reactions associated to rapid worm rejection [490] and sheep developing strong IgA response reducing female fertility [490] are considered as "resistant" to *T. circumcincta* based on their FEC. However, underlying mechanisms are different and certainly involve different genes. Hence, FEC shrinks the resistance repertoire and may reduce QTL detection power. Let's consider a simple case of two independent loci H and I respectively controlling hypersensitivity response and IgA response, for which two alleles are segregating H_a and I_a being associated to more favorable response than the H_b and I_b alleles. Co-segregation of H_a and I_b or H_b and I_a will result in low to moderate FEC as in one case, most of larvae will be expelled, while in the other case female worm fecundity will be maintained at low level. Therefore, the only observable differences will occur between the two H_a/I_a and H_b/I_b subsets, hence reducing the sample size for testing H and I locus effect. On the contrary, differentiating hypersensitive animals from IgA responders on the phenotypic scale would result in significant association with associated loci.

Most of the already published QTL analyses for resistance to GIN in sheep have considered Ig response [108, 128, 203, 454]. This trait was also considered in this work. However it is also a resultant of multiple previous steps. Other phenotypes could consist in measuring Th-2 cytokines expression or concentration within abomasal mucosa, in characterizing T-cells in the early steps of the acquired response within draining lymph nodes, or in quantifying amount of effector cells (eosinophils, mast cells) infiltrating the abomasal mucosa. This latter trait has already been surveyed in mice for resistance to *Heligmosomoide polygyrus* [358], while in sheep, Dominik *et al.* measured circulating eosinophils [128].

A major obstacle to the achievement of such an experiment are the associated costs, as sheep need to be sacrificed and phenotypes are far more expensive than FEC on both financial and time scales. Therefore these traits have not been measured in the BC flock. During the validation process, some samples have been taken to compare the mucosal infiltration by

eosinophils between the susceptibility groups. However, not enough time remained for analyzing these samplings.

9.4.2 How to consider gene interactions ?

Most of quantitative traits have a polygenic basis and resistance to GIN infection does not deviate from the rule [270]. This has been well illustrated by the number of QTL we found in this study and by the significant effect of the genomic breeding value estimated during the functional validation.

On the contrary, the QTL detection analyses we ran did not formally take this parameter into account. However, fitting a gEBV as a covariate while running analysis with the QTLMAP software did not change the resulting likelihood (data not shown). But from a theoretical standpoint, not considering putative interactions between QTL certainly provides a biased view of the genetic determinism of resistance to GIN.

Abundant literature has been produced about considering more than one QTL at a time while performing QTL analysis [259, 576, 182]. However, computational time exponentially increases with the number of considered QTL. In addition, the more QTL fitted in the model, the more statistical levels to be tested. Hence, the sample size needs to dramatically increase with the number of QTL fitted in the model.

One way to overcome this problem is to mix positional candidates with already accumulated functional knowledge from web databases (see 9). In the frame of this study, it is interesting that OAR14 appeared central in the immune-related pathways analyses whereas the two QTL found on this chromosome were inconsistently associated to pepsinogen concentration (LA, LDLA) or mucosal IgG concentration and packed-cell volume (LDLA). This discrepancy between the gene network analysis and the QTL mapping underlines well the potential advantage of adding the functional knowledge to the QTL mapping analysis.

Even if this approach is biased toward already existing knowledge, it should be systematically include into the QTL mapping process to better integrate the fact that we are having a look at a whole biological system. This approach would consist in performing a QTL analysis with no *a priori* on the data in order to unravel as much relevant regions as possible. Following this analysis, genes underlying the most plausible QTL regions would be fitted together into a gene network analysis. For every gene involved in a function related to the trait of interest, the likelihood of the corresponding QTL could be updated to higher values. Similarly, detected

QTL could also be gathered according to the metabolic pathways they are related to. Hence, QTL detection experiments would be dedicated to unravel significant pathways rather than chromosome segments. This would be a better approach for fitting the biological reality that each QTL region rarely works on its own but in close relationship with other regions inside a dedicated functional network.

9.4.3 Working on a biological system

9.4.3.1 Considering host-pathogen interactions

Resistance to GIN is a dynamic phenomenon. Indeed, the immune response is, by essence, adaptive. This is well underlined by the splitting of the response into the innate and acquired responses (see 1.3) and by the so-called "heterozygote advantage" observed at the *DRB1* locus, whose variability increases pathogen recognition repertoire and host resistance [490]. Further, GIN are also "masters of regulation" [331]. They are capable of polarizing the immune response, and they can hide themselves from recognition mechanisms [226, 516]. In turn, infection outcome is the sum of the many interactions between the parasite and its host. Hence, considering the only sheep side of the matter is expected to provide a truncated view of the system.

One major limitation is how to take these interactions into account. Firstly, time-series data are worth improving the knowledge gained from QTL analyses by accounting for dynamics [440]. To screen for differential candidate gene expression following infection by GIN, designs usually considered two time points, before and after infection [316, 265, 266]. But studies considering several time-points under infection are costly and thus scarce [327, 255]. One question immediately arises: *"when should the system be under study ?"*

The first 48 hours after infection should be relevant as they represent first interactions between the host and worms, and crucial steps like recognition condition the infection outcome. Additional time points are more difficult to set. One could focus on the nematode life stages as changes from one state to another might be associated to changes in epitopes and various sheep/nematode interactions. As it has been demonstrated that heritability of resistance to GIN increases with lambs age [566, 46, 45, 419], it could be interesting to perform such study in already challenged lambs to better target the adaptive genetic component of resistance to GIN.

Secondly, it could be interesting to simultaneously monitor changes in both the sheep and worms population. Not only unraveling QTL affecting resistance of sheep, it could give access to nematode genes involved in survival. Functional investigation on these regions could help

detecting functional effectors that help escaping from the host immune effectors thus bringing to light critical mechanisms of the immune response. In the same way, gene expression analysis performed in such system could help drawing network of genes involved during sheep/nematode interactions. Unfortunately, nematode genomes are poorly known and sequencing, if attempted, has not been completed yet [289], thus hampering genetic marker development and gene network study. In addition, nematodes undergo many physiological transformations during their parasitic life that might be independent from their host reaction, and that could interfere with the "true" sheep/worm interactions.

9.4.3.2 How to reduce variability of parasite lines ?

GIN, including *H. contortus* are known to have a huge genetic diversity [66]. Because of this diversity, it is difficult to properly standardize the infection challenge of sheep, so that non-genetic inter-sheep variation occurs.

To get rid of this effect it could be interesting to dispose of fixed line of *H. contortus*. At least an investigation on the genetic structure of the Humeau line as well as a regular monitoring of its genetic diversity could help controlling this point. Still, this matter may not be relevant as challenge is done using a 10,000 larvae infection dose. Provided the Humeau line has been maintained by using susceptible animals with diverse genetic background, selection pressure on this line should have been low. Therefore it can be anticipated that genetic diversity between infection doses should be low within an experiment, hence limiting between-sheep variations.

Whether findings of one experiment are affected or not by the batch of larvae remains unresolved. As well, between-labs comparison of findings might be affected by such genetic variation between considered lines.

9.4.3.3 Tissue models as alternative systems ?

To simplify the study of resistance to GIN, the system could be reduced to its basic components. In the frame of resistance to GIN, tissue models could be relevant. In a proteomic approach, Athanasiadou *et al.* considered an *in vitro* model of abomasal mucosa to investigate consequences of *T. circumcincta* challenge [23]. In the same way, Terefe *et al.* studied direct interactions between eosinophils and *H. contortus* L3 larvae [**?**, 508].

Even if these systems are efficient in simplifying the complexity of the problem, they need strong technical skills to be set up, especially *in vitro* culture of abomasal tissue. Beyond the

only material issue, their pertinence is questionable as they sweep the *in vivo* environment and might lead to spurious conclusions. For instance, the induction and regulation of the immune response is made of cascade and homeostatic control that may not be reproduced in *in vitro* conditions.

Successful application of such models have been reported in case of unicellular organism [61]. It is interesting to point out that even in presence of a far simpler system, strong computational and experimental developments were required to reach a complete understanding of the biological system.

9.4.3.4 Any role for the metagenome ?

The *B. thuringiensis* is a parasite of the *C. elegans* nematode and the study of their interactions has provided interesting insights about co-evolution theory [465, 466]. Within their hosts, GIN are certainly in close relationship with the host bacterial flora. Bacterial flora properties are largely unknown and metagenomics studies have flourished within recent years. Hence, it can be hypothesized that abomasal or gut flora (according to the parasite species) could alter parasites survival or could contribute to its elimination from the host.

Metagenomics studies have looked at modifications occurring during helminth infection in mice [543], pigs [569, 314] and cattle [313]. All experiments but the *O. ostertagi* infection in cattle resulted in altering the intestinal flora composition. These results deserve further investigations to explore any clinical consequences of these shifts. It could be interesting to investigate whether any differences in intestinal microbiota composition could alter nematode survival.

9.5 Perspectives on the practical implementation of genetic selection for resistance to GIN

The three years of this PhD project have been devoted to unravel one or some genes affecting resistance to GIN. One of the major achievement would have been to identify one gene. This quest for the Holy Grail has been partially achieved. Two QTL have been particularly investigated but no causative mutation can be proposed so far.

Interestingly, while this huge effort has been put on the identification of genes, a huge amount of literature had already been accumulated about the *MHC* locus (see 4.2). This gene has been known for a long time and its effect on resistance to GIN has been the most consistent across

studies on resistance to GIN. Despite the characterization of this locus, no successful selection of this region has been reported for resistance to GIN. This is mainly because ways of action are still far from understood. Further, it seems that variability at this locus confers a fitness advantage [488] so that selection for a fixed combination of alleles at this locus is not really indicated.

Therefore, it seems that unraveling genes only marks the beginning of another big amount of questions. In addition to the questions about the objective assessment of the candidates resistance (fully discussed in the literature review, see 3.2.2.2), remaining questions fall into the following categories:

1. What will be the consequences of the selection for resistance to GIN ?

2. How to implement genetic resistance on field ?

Considering the results of the QTL analyses and the respective functional investigations, what should be done with the identified QTL, with or without the causative mutation ?

9.5.1 It seems a limited number of regions exert a non negligible effect

Within the BC design, a limited number of QTL regions were identified. Among these, the QTL located on OAR12 appears to be a key player in resistance to *H. contortus*. After functional investigation, its effect seems to impact female worms fertility which is of particular interest in reducing pasture parasitic burden.

Though the identification of the underlying QTN has not been achieved yet and the region under investigation still spans a non-negligible interval of OAR12, hence hampering its use in selection. A complete understanding of the genetic determinism of the QTL effect is also required before launching any selection process in order to avoid any detrimental effect on production or susceptibility to other diseases. Indeed, improving resistance to GIN infection should result in a more efficient anti-helminth immune response. One way of improving the anti-helminth response could be achieved by a higher tendency to shift toward the Th-2 type response, hence hampering the Th-1 response. This question still needs to be addressed.

Knowledge about host resistance mechanisms could also be used to design combination of QTL to be selected for that could maximize the probability of success while minimizing the probability of parasite adaptation. This could be achieved by targeting mechanisms simultaneously preventing worm installation in its host while preventing its subsequent development from

one stage to another.

Not only providing a better understanding of the genetic architecture, accessing the underlying QTN should simplify the selection across breeds. Indeed, handling QTL in selection requires to control that the linkage phase between markers and the QTL holds from the breed it has been detected to other breeds in which it should be selected. Knowing the underlying gene, it is easy to check whether favorable or unfavorable alleles are segregating in the breed of interest.

In case, the QTL/QTN is not segregating in the breed of interest, introgression could be considered to introduce the most favorable alleles on resistance. For instance, some crosses between RMN and MBB sheep followed by successive back-crossing on RMN sheep could allow the introduction of MBB alleles within the RMN breed. This process is not easy to implement as QTL allele needs to be traced at each step of mating to ensure that it is transmitted to the next generation. This point tackles another related aspect of implementing selection for resistance to GIN. Instead of introgressing favorable alleles from the MBB breed within the RMN breed, a first easier step could the identification and elimination of susceptible alleles from the RMN population. This would contribute to reduce the eggs output by eliminating sheep with highest worm burden.

It seems a limited number of regions were of primary importance to explain differences of susceptibility between the MBB and RMN breeds. An interesting question is the number of regions to be considered for breeding purposes. Indeed, the more QTL involved, the more difficult it is to simultaneously select for the most favorable combination of alleles. In addition, it is difficult to fit a threshold on the QTL effect for considering it or not for breeding purpose. Weller et al. suggested to not include any QTL if its effect was less than 1% of the genetic variation [552].

Tuning the selection of resistant sheep through the selection of particular combinations of regions exerting preponderant effects should lead to a more rapid genetic progress while preventing the occurrence of any detrimental effects. Candidates to genetic evaluation could be genotyped for a restricted set of causative markers fitted on a small dedicated SNP chip, hence reducing the associated costs.

9.5.2 What about a polygenic approach ?

In addition to the limited number of regions that showed a non-negligible effect, many other QTL have been unraveled in this study. As well, Kemper et al. highlighted the polygenic nature

of resistance to GIN infection [270].

Taking this polygenic nature into account, Australia and New-Zealand have definitely focused on a robust black box approach of whole genome selection without knowledge about mechanisms [263]. This has the particular advantage of selecting the genome as a whole, hence avoiding to target a particular mechanisms of resistance. In plants, the breeding of more resistant lines have relied on a limited pool of genes [556, 418, 480]. However, each plant resistance gene usually targets a corresponding pathogen gene, which in case of mutation overwhelm the plant resistance. This has been described in wheat, tomato, pepper or potato for instance [480, 418]. Resistant parasites thus become preponderant even if their fitness is reduced as they are the only one able to develop [480]. Due to the limited pool of genes available, this has lead to a race between pathogens and breeders who now reach the point where few genes remain for breeding purposes [418, 480].

On the contrary, if many genes of weak to moderate effects are selected together, they all apply simultaneously a weak selection pressure on multiple components of the worms rejection. This multiplicity of targets should prevent the apparition of worms capable of adapting to their hosts [485]. Kemper et al. failed to find evidence of parasite adaptation to sheep selected for resistance [269]. However, *H. contortus* larvae able to settle and live in MBB sheep have been realized under controlled experimental conditions (A. Blanchard-Letort, personal communication). Whether this adaptation impacts the worm fitness is still under investigation.

Interestingly, insights from plants tend to prove that *"there is not a single genetic basis for durable resistance"* (Johnson, cited in [480]. Indeed, the recessive gene *mlo* for powdery mildew resistance or *Rgp1* for resistance to stem rust in barley are associated to multiple isolates and have provided durable resistance over decades of use in crop breeding [480].

Selection for the genome as a whole is of particular interest as no QTL knowledge is needed. Whether this should be implemented through a classical selection scheme or rather through genomic selection approach depend on the sheep population structure. In France, sheep populations particularly demanding of genetic progress on resistance to GIN are meat breeds and the dairy breeds of Pays-Basque (Manech tête noire, Manech tête rousse and Basco-béarnais). In both cases, the constitution of a reference population is greatly hampered by the size of the breed population. As the Basque dairy breeds are in close genetic relationships, a reference set might be built using the three breeds. The genomic selection of meat sheep breeds is highly questionable. Indeed, breeds should also be gathered together into a single reference set to

achieve the 1,000 animals threshold [188]. Even if this reference population could be built, a denser SNP chip will be required to obtain accurate estimates of gEBV [103].

Therefore, for the French problematic it seems classical selection approach is the best suited option. In case, some causative mutations could be identified, candidate genotypes could be performed and integrated in the breeding value estimation.

9.5.3 Thinking "integrative management": practical implementation and hurdles on the way

9.5.3.1 Farmer is *THE* key player

The main focus of this work is to find solutions to the problem of anthelmintic resistances. Developing new management strategies less reliant on anthelmintics appears of primary importance to preserve the efficacy of anthelmintic classes that have not been endangered by worms resistances.

Proposing new methods of GIN management have been a dynamic field of research but, so far, no optimal alternative strategy has been developed. This failure does not come from a lack of imagination from researchers, but is rather in strong relationship with the farmers acceptance of these methods. Indeed, farmers will not only consider the effectiveness, but also cost and ease of applying the strategy. So that the "sustainability" of an option will be ranked at a lower order priority [34]. This was also reported by Kenyon *et al.* [273] that stated the targetet selected treatment indicator should be "quick and simple to use, cost-effective, easily learned and allow treatment decisions to be made 'sheep-side"'. Complementary findings of Cabaret *et al.* from a survey of 16 conventional and organic farms in France showed that human aspects were determinant in the implementation of integrated nematode management [78]. He also showed how social aspects determined the decision of treating or not animals [78].

Van Wyk also detailed several aspects of the difficulty to integrate sustainable alternative ways of control of nematodes [523]. He pointed out the fact that the problem of arising resistance to anthelmintics is not always well understood whereas integrated management of nematodes increases the complexity of treatment [523].It is hence of primary importance that strong communication and education is made available to both farmers and their advisors to ensure that the new strategies are fully understood as well as the related benefits [272].

Very few studies aimed at quantifying the costs of implementing such TST [44]. Some results obtained by Mahieu *et al.* [329] (cited in [44]) for creole goats, pointed out that the

implementation of the FAMACHA system was associated to an economic loss, that was balanced with the reduced selection of anthelmintic resistance and the reduction of haemonchosis.

9.5.3.2 The unique farmers-flock couple and veterinarians

In the frame of a study aiming at developing new alternative management of GIN in the French Pays Basque, a panel of farms have been chosen and interviews about sheep and GIN management have been performed (P. Jacquiet, personal communication). While performing these interviews, it was striking that every farm had its own working rules: each flock differed in size, the available pasture surface differed and applied anthelmintics also differed. Some breeders try to change their treatment, other try to implement targeted selective treatment while some apply the same treatment procedure year after year.

Therefore, thinking an integrative management is local issue. Various alternative ways of GIN control have been proposed but they are not all suitable for every farm. For instance, typical French Pays Basque pastures are tiny and sparsed over a huge area which make it impossible to implement pasture management: due to the limited surface, sheep usually have to move from one pasture to another on the same day.

Further, some of the flock can be brought to summer pastures located in Pyrenees. These mountain pastures are then grazed by several flocks together, hence facilitating the transfer of resistant worms from one flock to another.

To this regard, veterinarians should reconquer the parasitology field. Indeed, they are the main actors of animal health and they dispose of a broaden view on how the farm works. Therefore, they should not only sell anthelmintics, but also design an integrated pest control programme that would consider the whole farm constraints and opportunities. Breeders are the center of a complex ecosystem of counselors (AI technicians, food sellers, veterinarians, milk technicians ...). Each of these actors usually provide more or less enlightened advice on animal health management. In my opinion, GIN management represents a neglected aspect of veterinary medicine, that is usually solved by a blinded application of anthelmintics with no additional foreseeing. The commercial offer and strong incomes anthelmintics represent, is also questionable for their sustainable use.

9.5.3.3 The need for passive approach

Anthelmintics are cheap, easy to use and still efficient in most temperate area (except for benzimidazoles). As such, they still remain the most important option to achieve nematode control. In a way, they constitute a blessing for breeders. An efficient drenching has quick seenable consequences on sheep health. For dairy breeders, drenching is usually concomittant of an increase in milk production. Within a week, he can materialize the benefits of using anthelmintics.

On the contrary, alternative strategies often require changing the flock management (which might be simply difficult to admit for the breeder). They are also often synonymous of an increased work load (pasture management, sorting the animals to be drenched). To these inconvenient inputs, alternative strategies usually generate few outputs if any as, in regions where anthelmintics are still efficient, they only prevent the appearance of a putative long-term issue. So that no benefits are perceived by breeders.

This paradox is contained in one of the simplest approach of reasoned GIN management. To limit the use of anthelmintics and prevent appearance of resistance, it is possible to perform coproscopic examination to assess worm burden. In case, no treatment is required, the only associated cost is the 20€/ coproscopy. In case a treatment is required, breeder will pay both coproscopy (20€/ treatment) and treatment (1 to 2€/ sheep). In case, he decides to treat every sheep, he will only pay the treatment which represent at most 2€/ sheep ...

9.5.3.4 Potential for the genetic approach

The breeding of more resistant animals seems to be an ideal alternative strategy. Indeed, at farm level it means no additional work load for an efficient implementation. Thus, from a breeder perspective, it offers a passive improvement of the sheep potential, that makes genetics one of the most appropriate complementary strategy of GIN control. In addition, it contributes to the removal of worms from pasture which is a non-negligible epidemiological feature.

The question of extra-costs still remains. Indeed, under a classical selection scheme, selection candidates need to be challenged, which hence generate some extra-costs : preparation of the infection dosis, coproscopy. Assuming this should be implemented on a national scale or even European scale, prices for such evaluation should go down. In the end, a 50 €extra-charge per ram should not be exceeded which hence represent around 10% increase of its price. This is relatively cheap in comparison to the associated long-term benefits.

Further, the breeding of more resistant animals cannot be a unique strategy, as it aims

at a long-term benefit which is in strong opposition to the short-term economic reality of the breeder's income. Indeed, the response to selection, R, depends on the intensity of selection in each sex, i_m and i_f for males and females respectively, the generation interval in males and females (L_m and L_f respectively), the heritability of the trait, h^2 and the phenotypic standard deviation, σ_p [151] as:

$$R = \frac{\sum i}{\sum L} h^2 \sigma_p$$

Let us consider generation intervals of the Manech dairy breed $\sum L = 3.02+3.89+4.48+3.5 = 14.89$ [104] and the following selection intensities (J. M. Astruc, personal communication) for sires of males ($i = 1.03$), sires of females ($i = 0.56$), dams of males ($i = 0.92$) and dams of females ($i = 0.16$). This gives $\sum i = 2.67$. In the end, a $0.05\sigma_p$ reduction of FEC output should be achieved at each generation. This strategy hence needs to be implemented at the population level, so that it will benefit a whole region on a long-term time scale. In a situation of endemic resistance to anthelminthics like in French Pays Basque, resistant sheep would both contribute to a farming less reliant on anthelmintics and to reduce the frequency of resistant populations by limiting the amplification of resistant worms populations. Other alternative strategies should be applied as well to handle the "emergency" of the situation at the farm-scale, taking into account every particular factor. These strategies should be managed by a limited number of actors to further both strong communication without discrepancies in the delivered message and a strong follow-up.

Part VI

Conclusion

La sélection d'ovins disposant d'une meilleure résistance aux strongles gastro-intestinaux semble être une stratégie de contrôle envisageable et *a priori* durable. Cependant, la mise en place d'une telle sélection réclame une meilleure compréhension des interactions entre les ovins et les strongles. Il a été démontré que la race Martinik Black-belly est plus résistante à *H. contortus* que la race Romane et des approches d'immuno-pathologie comparative ont mis en évidence une réponse immunitaire plus forte chez les Martinik Black-belly. Dans le cadre de ce projet de thèse, nous nous proposions de disséquer le déterminisme génétique de ces différences raciales. Une approche en deux temps a été mise en place, visant premièrement, à identifier les régions du génome impliquées dans la variation de sensibilité à *H. contortus* et ensuite, à mettre en place une validation fonctionnelle des régions d'intérêt significatif.

Parmi les régions QTL identifiées, deux régions localisées sur OAR12 et OAR21 présentaient un intérêt particulier et ont donc fait l'objet d'une étude approfondie.

D'une part, une région de 10 Mbp portée par OAR12 a été associée communément à l'intensité d'excrétion d'oeufs en première et deuxième infestation. Le QTL associé à l'intensité d'excrétion d'oeufs en première infestation présentait l'un des effets les plus forts de $0.19\sigma_p$ en analyse de liaison, tandis que la différence d'effets entre les allèles les plus extrêmes identifiés dans l'analyse d'association dépassait $1\sigma_p$. Par ailleurs, une approche de génétique des populations entre les races MBB et RMN a révélé une trace de sélection fixée dans la population MBB, et localisée dans la région QTL d'intérêt. En supposant que cette trace de sélection était responsable de l'effet QTL détecté, une population de validation a été créée, en croisant des animaux back-cross sur la base de leur génotype à cette trace de sélection. Peu de différences ont été trouvées entre les groupes constituées. Au contraire, en utilisant l'haplotype de 4 SNPs maximisant la vraisemblance en analyse d'association pour trier les animaux croisés des différences significatives sont démontrées entre les prédits résistants et sensibles. Un effet significatif sur la fertilité des vers femelles associé à une imprégnation Th-2 plus forte de la muqueuse abomasale chez les animaux prédits résistants a en effet été mis en évidence. Par ailleurs, aucun gène candidat fonctionnel évident n'a pu être identifié.

D'autre part, une région d'OAR21 a été associé de manière significative à l'intensité d'oeufs excrétés considérée comme mesure répétée et à la variation de concentration en pepsinogène, marqueur biochimique d'infestation utilisé en médecine vétérinaire pour évaluer l'état d'infestation d'un troupeau. Le locus *PGA5* qui code pour le pepsinogène a rapidement été identifié comme un candidat fonctionnel plausible. Le séquençage des introns et des exons dans des groupes

génétiquement prédits comme “haut” et “bas” producteurs de pepsinogène a révélé neuf marqueurs introniques spécifiques de chaque groupe. De manière intéressante, des corrélations favorables ont été estimées entre la variation en pepsinogène plasmatique et l'intensité d'excrétion d'oeufs dans la population back-cross et la population de croisés BCxBC. De plus, les groupes allèliques constitués ont montré des différences significatives dans l'installation des vers. Le test des polymorphismes identifiés sur le gène *PGA5* est en cours, mais nous pensons qu'il serait judicieux de tester aussi des polymorphismes situés à proximité des loci *CD5* et *CD6* impliqués dans la régulation des lymphocytes T [398].

En plus de ces deux regions QTL, OAR14 est apparu central dans les réseaux métaboliques liés à la réponse immunitaire. De plus, d'autres équipes de recherche ont également identifié cette région comme ayant un effet prépondérant dans la résistance à l'infestation par des strongles gastro-intestinaux. Cette région pourrait être sujet d'une future étude.

Ce travail de thèse a apporté quelques pièces complémentaire au puzzle de la résistance génétique aux strongles gastro-intestinaux. Désormais, un effort de recherche supplémentaire est nécessaire pour disséquer plus amont ces régions pour obtenir une meilleure vision de l'architecture génétique de la résistance à l'infestation par des strongles gastro-intestinaux. Des approches de reséquençage total ou d'études d'expression par séquençage d'ARN sont des options de recherche valables. Cela permettra non seulement d'anticiper de quelconques effets indésirables sur les moutons sélectionnés, mais aussi d'identifier la(es) mutation(s) causale(s) qui faciliteraient la sélection dans plusieurs races.

De nombreuses autres questions restent en suspend sur l'utilisation de gènes de résistance aux strongles gastro-intestinaux: la présence des polymorphismes de résistance dans la race d'intérêt, le taux de sélection à utiliser au sein d'un troupeau pour bénéficier d' avantages épidémiologique et économique suffisants ou bien encore la gestion intégrée de la génétique et du traitement anthelminthique. De plus, les questions d'importance sur l'apparition éventuelle d'effets indésirables sur la sensibilité à d'autres pathogènes ou la réduction potentielle des capacités de production reste inconnu.

Le but principal de ce travail est de trouver des solutions complémentaires aux anthelminthiques. Des efforts de développement de stratégies intégratives de gestion du parasitisme qui reposeraient sur une utilisation moindre des anthelminthiques sont nécessaires pour préserver l'efficacité des classes d'anthelminthiques les moins touchées par les résistances des vers. Les anthelminthiques sont bon marché, faciles à utiliser et encore efficace dans les régions tempérées (à l'exception

des benzimidazoles). A ce titre, les anthelminthiques restent l'option de controle des nématodes la plus importante. Le développement de nouvelles méthodes de gestion des strongles gastro-intestinaux a constituté un champ de recherche dynamique, mais aucune stratégie optimale n'a pu être proposée jusqu'à présent. La sélection d'animaux plus résistants semble être un moyen complémentaire satisfaisant. En effet, sa mise en place ne requiert ni un surcoût ni une charge de travail supplémentaire pour l'éleveur. Ainsi, du point de vue de l'éleveur, cette alternative génétique représente une amélioration passive du potentiel du troupeau, qui place la génétique comme l'un des moyens de gestion des nématodes les plus appropriés.

Cependant, la génétique doit être considérée comme un moyen complémentaire à intégrer dans une stratégie de contrôle intégrée qui doit agir sur les points critiques déjà connus (traitement administré correctement, vermifugation de quarantaine, amélioration génétique et gestion de paturages) qui envisagent aussi les propriétés épidémiologiques des infestations par les strongles gastro-intestinaux (ciblage des populations les plus à risques, adaptation aux conditions météorologiques locales). La sélection génétique d'hôtes résistants peut améliorer la population ovine à l'échelle d'un pays, mais la gestion durable des anthelminthiques passe par des facteurs propres à chaque élevage. Chaque couple éleveurs-élevage constitue un microcosme unique qui doit être abordé au cas par cas pour la mise en place de solutions efficaces et pertinentes.

The breeding of sheep with a better ability to cope with gastro-intestinal nematodes has been devised as a feasible and potentially sustainable alternative strategy for worm control. Still, questions remain to be addressed before launching such selection and a more precise understanding of the sheep-nematode interactions is required. The MBB breed has been demonstrated to be more resistant to *H. contortus* infection than its RMN counterpart, and patho-physiological approaches have shown that the immune response of MBB sheep was stronger. In this PhD project, we proposed to mine the genetic architecture of the observed differences between the Martinik Black-belly and the Romane breeds. A two-step approach was implemented to firstly identify regions of the genome particularly involved in resistance to an experimental challenge by *H. contortus* and secondly, functionally investigate some outstanding positional candidates. QTL mapping analyses highlighted several regions whose variability could be worth further studies and eventually selection.

Among these, two QTL regions on OAR12 and OAR21 were thought of particular interest and deserved additional investigations during this PhD project.

On one hand, OAR12 carried a 10 Mbp region that was common to two QTL affecting FEC at first and second infection respectively. The QTL affecting FEC in naïve lambs exhibited a $0.19\sigma_p$ effect, while the most extreme allelic effects of the association analysis were slightly above $1\sigma_p$. In addition, a population genetics approach highlighted a selection signature occurring in the resistant MBB breed, within this QTL region. This region was thought to contribute to the QTL likelihood and a validation population was build by crossing BC sheep according to their sweep genotype. Few differences were found between sweep-based allelic classes. On the contrary, a 4-SNP haplotype located at 56 Mbp was efficient at tagging true resistant and susceptible animals. It showed an effect on female worm fertility that was concomitant of a stronger Th-2 cytokinic environment in carriers of one MBB allele associated to resistance. In addition, no obvious candidate gene have been identified so far.

On the other hand, OAR21 was significantly associated to FEC considered as a repeated measure and to pepsinogen concentration which has been used in veterinary medicine to monitor infection burden at the flock scale. The *PGA5* locus which codes for the pepsinogen was rapidly identified as a probable functional candidate. Sequencing of exons and introns in genetically predicted "high" and "low responders" revealed nine intronic markers specific of each allelic group. Interestingly, favorable correlations were found between pepsinogen variation and FEC in both BC and BCxBC populations and QTL allelic groups showed significant differences in worm burden. Polymorphisms identified in the *PGA5* locus are currently being tested. We think that an investigation of segregating polymorphisms in the two other positional candidates (*CD5* and *CD6* loci) that are known to act as T-cell regulators and that are located in the vicinity of *PGA5* should be performed.

In addition to these two QTL regions, OAR14 emerged from a gene network analysis as a central region involved in immune-related pathways. Further, other independent research teams also found this chromosome among top players in resistance to GIN infection. This region could also be under investigation for future projects.

This work brought a little piece to the puzzle of resistance to GIN infection. Now, additional efforts are needed to dissect further this limited region of the genome to better understand the genetic architecture of resistance to GIN infection. Whole-genome resequencing and RNA-seq approaches would be valuable research track to be explored.This will not only help anticipating any detrimental effect on sheep selected for a particular QTL allele but it should also lead to the causative mutation identification that would simplify selection across breeds.

Many other questions remains on the use of genes affecting resistance to GIN about the segregation of interesting polymorphisms in the breed of interest, the selection rate to be applied within flocks to benefit both epidemiological and economical benefits or the integrated management of genetics and other alternative strategies. Further, the main questions about occurrence of detrimental effects on the susceptibility to other diseases or the putative reduction in production abilities remain unresolved.

The main focus of this work is to find solutions to the problem of anthelmintic resistances. Efforts should be put on developing breeding strategies less reliant on anthelmintics in order to preserve the efficacy of anthelmintic classes that have not been endangered by worms resistances. Anthelmintics are cheap, easy to use and still efficient in temperate area (except for benzimidazoles). As such, they still remain the most important option to achieve nematode control. Proposing new methods of GIN management have been a dynamic field of research but, so far, no optimal alternative strategy has been developed. The breeding of more resistant animals seems to be a good alternative strategy. Indeed, at farm level it means neither extra-costs for sheep breeders nor additional working load for being efficient. Thus, from a breeder perspective, it offers a passive improvement of the sheep potential, that makes genetics one of the most appropriate complementary strategy of GIN control.

In addition, genetics should be considered as one alternative among others. GIN infection management should be thought of in an integrated fashion by simultaneously acting on every identified critical points (proper anthelmintic treatment, quarantine drenching, host genetics improvement, grazing management) and bearing in mind epidemiological features (targeting susceptible populations, adapting to local weather conditions). Genetic selection can help improving sheep population at a country scale, but the sustainable management of anthelmintic drugs is a farm-dependent issue that cannot be resolved by general rules. Each farm and its associated-management are a unique environment that needs to be considered in a case-by-case approach.

Bibliography

[1] *Dict. Medicaments Veterinaires Et Produits Sante Animale 2011.* Educa Books, 2011.

[2] K. Abbott, M.A. Taylor, and L. A. Stubbings. *Sustainable Worm Control Strategies for Sheep. A technical manual for Veterinary Surgeons and Advisors (2nd ed.), SCOPS.* Malvern, 2007.

[3] J Aerts, H J Megens, T Veenendaal, I Ovcharenko, R Crooijmans, L Gordon, L Stubbs, and M Groenen. Extent of linkage disequilibrium in chicken. *Cytogenetic and Genome Research*, 117(1-4):338–345, 2007. PMID: 17675876.

[4] M.S Akhtar, Zafar Iqbal, M.N Khan, and Muhammad Lateef. Anthelmintic activity of medicinal plants with particular reference to their use in animals in the Indo–Pakistan subcontinent. *Small Ruminant Research*, 38(2):99–107, October 2000.

[5] F. Alba-Hurtado, E. Romero-Escobedo, M.A. Muñoz-Guzmán, G. Torres-Hernández, and C.M. Becerril-Pérez. Comparison of parasitological and productive traits of criollo lambs native to the central mexican plateau and suffolk lambs experimentally infected with haemonchus contortus. *Veterinary Parasitology*, 172(3-4):277–282, September 2010.

[6] G A Albers, G D Gray, L R Piper, J S Barker, L F Le Jambre, and I A Barger. The genetics of resistance and resilience to haemonchus contortus infection in young merino sheep. *International journal for parasitology*, 17(7):1355–1363, October 1987. PMID: 3429127.

[7] F.A. Almeida, K.C.O.D. Garcia, P.R. Torgerson, and A.F.T. Amarante. Multiple resistance to anthelmintics by haemonchus contortus and trichostrongylus colubriformis in sheep in brazil. *Parasitology International*, 59(4):622–625, December 2010.

[8] A. J. Amaral, H. J. Megens, R. P. Crooijmans, H. C. Heuven, and M. A. Groenen. Linkage disequilibrium decay and haplotype block structure in the pig. *Genetics*, 179(1):569–79, 2008.

[9] A F Amarante, T M Craig, W S Ramsey, S K Davis, and F W Bazer. Nematode burdens and cellular responses in the abomasal mucosa and blood of florida native, rambouillet and crossbreed lambs. *Veterinary Parasitology*, 80(4):311–324, January 1999. PMID: 9950337.

[10] A F Amarante, T M Craig, W S Ramsey, N M El-Sayed, A Y Desouki, and F W Bazer. Comparison of naturally acquired parasite burdens among florida native, rambouillet and crossbreed ewes. *Veterinary Parasitology*, 85(1):61–69, August 1999. PMID: 10447193.

[11] A F T Amarante, P A Bricarello, J F Huntley, L P Mazzolin, and J C Gomes. Relationship of abomasal histology and parasite-specific immunoglobulin a with the resistance to haemonchus contortus infection in three breeds of sheep. *Veterinary Parasitology*, 128(1-2):99–107, March 2005. PMID: 15725538.

[12] A F T Amarante, P A Bricarello, R A Rocha, and S M Gennari. Resistance of santa ines, suffolk and ile de france sheep to naturally acquired gastrointestinal nematode infections. *Veterinary Parasitology*, 120(1-2):91–106, February 2004. PMID: 15019147.

[13] A F T Amarante, I Susin, R A Rocha, M B Silva, C Q Mendes, and A V Pires. Resistance of santa ines and crossbred ewes to naturally acquired gastrointestinal nematode infections. *Veterinary Parasitology*, 165(3-4):273–280, November 2009. PMID: 19656629.

[14] Leif Andersson and Michel Georges. Domestic-animal genomics: deciphering the genetics of complex traits. *Nature reviews. Genetics*, 5(3):202–212, 2004. PMID: 14970822.

[15] Wilfried Andlauer and Peter Fürst. Nutraceuticals: a piece of history, present status and outlook. *Food Research International*, 35(2-3):171–176, January 2002.

[16] Nicholas Andronicos, Peter Hunt, and Ross Windon. Expression of genes in gastrointestinal and lymphatic tissues during parasite infection in sheep genetically resistant or susceptible to trichostrongylus colubriformis and haemonchus contortus. *International Journal for Parasitology*, 40(4):417–429, March 2010. PMID: 19825375.

[17] Robert M Anthony, Laura I Rutitzky, Jr Urban, Joseph F, Miguel J Stadecker, and William C Gause. Protective immune mechanisms in helminth infection. *Nature reviews. Immunology*, 7(12):975–987, 2007. PMID: 18007680.

[18] A L Archibald, N E Cockett, B P Dalrymple, T Faraut, J W Kijas, J F Maddox, J C McEwan, V Hutton Oddy, H W Raadsma, C Wade, J Wang, W Wang, and X Xun. The sheep genome reference sequence: a work in progress. *Animal Genetics*, 41(5):449–453, October 2010. PMID: 20809919.

[19] K. G. Ardlie, L. Kruglyak, and M. Seielstad. Patterns of linkage disequilibrium in the human genome. *Nat Rev Genet*, 3(4):299–309, 2002.

[20] R Artho, M Schnyder, L Kohler, P R Torgerson, and H Hertzberg. Avermectin-resistance in gastrointestinal nematodes of boer goats and dorper sheep in switzerland. *Veterinary Parasitology*, 144(1-2):68–73, March 2007. PMID: 17088023.

[21] D Artis and R K Grencis. The intestinal epithelium: sensors to effectors in nematode infection. *Mucosal immunology*, 1(4):252–264, July 2008. PMID: 19079187.

[22] William Astle and David J. Balding. Population structure and cryptic relatedness in genetic association studies. *Statistical Science*, 24(4):451–471, November 2009.

[23] Spiridoula Athanasiadou, Alan Pemberton, Frank Jackson, Neil Inglis, Hugh R P Miller, Frank Thévenod, Annie Mackellar, and John F Huntley. Proteomic approach to identify candidate effector molecules during the in vitro immune exclusion of infective teladorsagia circumcincta in the abomasum of sheep. *Veterinary research*, 39(6):58, December 2008. PMID: 18715541.

[24] G. Aumont, L. Gruner, and G. Hostache. Comparison of the resistance to sympatric and allopatric isolates of haemonchus contortus of black belly sheep in guadeloupe (FWI) and of INRA 401 sheep in france. *Vet Parasitol*, 116(2):139–50, 2003.

[25] R. F. E Axford. *Breeding for disease resistance in farm animals*. CABI Pub., Wallingford, Oxon, UK; New York, 2000.

[26] Yvonne M Badke, Ronald O Bates, Catherine W Ernst, Clint Schwab, and Juan P Steibel. Estimation of linkage disequilibrium in four US pig breeds. *BMC Genomics*, 13:24, 2012. PMID: 22252454.

[27] Seyeon Bae, Hyemin Kim, Naeun Lee, Cheolhee Won, Hang-Rae Kim, Young-il Hwang, Yeong Wook Song, Jae Seung Kang, and Wang Jae Lee. α-enolase expressed on the surfaces of monocytes and macrophages induces robust synovial inflammation in rheumatoid arthritis. *Journal of immunology (Baltimore, Md.: 1950)*, 189(1):365–372, July 2012. PMID: 22623332.

[28] R.L. Baker, S Nagda, S.L. Rodriguez-Zax, B R Southey, J O Audho, E.O. Adula, and W. Thorpe. Resistance and resilience to gastro-intestinal nematode parasites and relationships with productivity of red maasai, dorper and red maasai x dorper crossbred lambs in the sub-humid tropics. *Animal Science*, 76:119–136, 2003.

[29] D. J. Balding. A tutorial on statistical methods for population association studies. *Nat Rev Genet*, 7(10):781–91, 2006.

[30] A Balic, V M Bowles, and E N Meeusen. Cellular profiles in the abomasal mucosa and lymph node during primary infection with haemonchus contortus in sheep. *Veterinary immunology and immunopathology*, 75(1-2):109–120, June 2000. PMID: 10889303.

[31] A Balic, V M Bowles, and E N T Meeusen. Mechanisms of immunity to haemonchus contortus infection in sheep. *Parasite immunology*, 24(1):39–46, January 2002. PMID: 11856445.

[32] A Balic, C P Cunningham, and E N T Meeusen. Eosinophil interactions with haemonchus contortus larvae in the ovine gastrointestinal tract. *Parasite immunology*, 28(3):107–115, March 2006. PMID: 16441509.

[33] I.A. Barger. Genetic resistance of hosts and its influence on epidemiology. *Veterinary Parasitology*, 32(1):21–35, July 1989.

[34] I.A. Barger. Prospects for integration of novel parasite control options into grazing systems. *International Journal for Parasitology*, 26(8-9):1001–1007, August 1996.

[35] Ian Barger. Control by management. *Veterinary Parasitology*, 72(3-4):493–506, November 1997.

[36] E H Barnes, R J Dobson, and I A Barger. Worm control and anthelmintic resistance: adventures with a model. *Parasitology today (Personal ed.)*, 11(2):56–63, February 1995. PMID: 15275374.

[37] N.J. Barton. Development of anthelmintic resistance in nematodes from sheep in australia subjected to different treatment frequencies. *International Journal for Parasitology*, 13(2):125–132, April 1983.

[38] K J Beh, D J Hulme, M J Callaghan, Z Leish, I Lenane, R G Windon, and J F Maddox. A genome scan for quantitative trait loci affecting resistance to trichostrongylus colubriformis in sheep. *Animal Genetics*, 33(2):97–106, April 2002. PMID: 12047222.

[39] M Benavides, T Weimer, M Borba, M Berne, and A Sacco. Association between microsatellite markers of sheep chromosome 5 and faecal egg counts. *Small Ruminant Research*, 46:97–105, November 2002.

[40] M.V. Benavides, T.A. Weimer, M.F.S. Borba, M.E.A. Berne, and A.M.S. Sacco. Genetic analyses of polymorphisms on ovine chromosomes 5 and 20 and their effect on resistance to internal parasites. *Small Ruminant Research*, 83(1-3):67–73, May 2009.

[41] B. O. Bengtsson and G. Thomson. Measuring the strength of associations between HLA antigens and diseases. *Tissue Antigens*, 18(5):356–63, 1981.

[42] D. Beraldi, A. F. McRae, J. Gratten, J. G. Pilkington, J. Slate, P. M. Visscher, and J. M. Pemberton. Quantitative trait loci (QTL) mapping of resistance to strongyles and coccidia in the free-living soay sheep (Ovis aries). *Int J Parasitol*, 37(1):121–9, 2007.

[43] B. Besier. New anthelmintics for livestock: the time is right. *Trends Parasitol*, 23(1):21–4, 2007.

[44] R B Besier. Refugia-based strategies for sustainable worm control: Factors affecting the acceptability to sheep and goat owners. *Veterinary Parasitology*, November 2011. PMID: 22197747.

[45] S. C. Bishop, K. Bairden, Q. A. McKellar, M. Park, and M. J. Stear. Genetic parameters for faecal egg count following mixed, natural, predominantly ostertagia circumcincta infection and relationships with live weight in young lambs. *Animal Science*, 63(03):423–428, 1996.

[46] S. C. Bishop, F. Jackson, R. L. Coop, and M. J. Stear. Genetic parameters for resistance to nematode infections in texel lambs and their utility in breeding programmes. *Animal*, 78:185–194, 2004.

[47] S C Bishop and M J Stear. Modeling of host genetics and resistance to infectious diseases: understanding and controlling nematode infections. *Veterinary Parasitology*, 115(2):147–166, July 2003. PMID: 12878420.

[48] S.C. Bishop, J. Chesnais, and M. J. Stear. Breeding for resistance to disease: issues and opportunities. Montpellier, 2002.

[49] S.C. Bishop and C.A. Morris. Genetics of disease resistance in sheep and goats. *Small Ruminant Research*, 70:48–59, June 2007.

[50] Stephen C Bishop and John A Woolliams. On the genetic interpretation of disease data. *PloS one*, 5(1):e8940, 2010. PMID: 20126627.

[51] S A Bisset and C A Morris. Feasibility and implications of breeding sheep for resilience to nematode challenge. *International journal for parasitology*, 26(8-9):857–868, September 1996. PMID: 8923135.

[52] S.A. Bisset, A. Vlassoff, C.J. West, and L. Morrison. Epidemiology of nematodosis in romney lambs selectively bred for resistance or susceptibility to nematode infection. *Veterinary Parasitology*, 70(4):255–269, July 1997.

[53] H Bjørn, J Monrad, A A Kassuku, and P Nansen. Resistance to benzimidazoles in haemonchus contortus of sheep in tanzania. *Acta Tropica*, 48(1):59–67, October 1990. PMID: 1980804.

[54] H Bjørn, J Monrad, and P Nansen. Anthelmintic resistance in nematode parasites of sheep in denmark with special emphasis on levamisole resistance in ostertagia circumcincta. *Acta Veterinaria Scandinavica*, 32(2):145–154, 1991. PMID: 1803927.

[55] N. Blake and G. Coles. Flock cull due to anthelmintic-resistant nematodes. *Veterinary Record*, 161(1):36–36, July 2007.

[56] Carine Blanchard and Marc E Rothenberg. Biology of the eosinophil. *Advances in immunology*, 101:81–121, 2009. PMID: 19231593.

[57] J.H. Boersema and V.S. Pandey. Anthelmintic resistance of trichostrongylids in sheep in the highveld of zimbabwe. *Veterinary Parasitology*, 68(4):383–388, March 1997.

[58] D Boichard, S Fritz, M. N. Rossignol, M.Y. Boscher, A. Malafosse, and J. J. Colleau. Implementation of marker-assisted selection in french dairy cattle. Montpellier, France, 2002.

[59] S. Boitard, C. Schlotterer, and A. Futschik. Detecting selective sweeps: a new approach based on hidden markov models. *Genetics*, 181(4):1567–78, 2009.

[60] Maxime Bonhomme, Claude Chevalet, Bertrand Servin, Simon Boitard, Jihad Abdallah, Sarah Blott, and Magali Sancristobal. Detecting selection in population trees: the lewontin and krakauer test extended. *Genetics*, 186(1):241–262, September 2010. PMID: 20855576.

[61] Richard Bonneau, Marc T Facciotti, David J Reiss, Amy K Schmid, Min Pan, Amardeep Kaur, Vesteinn Thorsson, Paul Shannon, Michael H Johnson, J Christopher Bare, William Longabaugh, Madhavi Vuthoori, Kenia Whitehead, Aviv Madar, Lena Suzuki, Tetsuya Mori, Dong-Eun Chang, Jocelyne Diruggiero, Carl H Johnson, Leroy Hood, and Nitin S Baliga. A predictive model for transcriptional control of physiology in a free living cell. *Cell*, 131(7):1354–1365, December 2007. PMID: 18160043.

[62] F H Borgsteede, J J Pekelder, D P Dercksen, J Sol, P Vellema, C P Gaasenbeek, and J N van der Linden. A survey of anthelmintic resistance in nematodes of sheep in the netherlands. *The Veterinary Quarterly*, 19(4):167–172, November 1997. PMID: 9413114.

[63] Nathan J Bott, Bronwyn E Campbell, Ian Beveridge, Neil B Chilton, Dianne Rees, Peter W Hunt, and Robin B Gasser. A combined microscopic-molecular method for the diagnosis of strongylid infections in sheep. *International journal for parasitology*, 39(11):1277–1287, September 2009. PMID: 19328802.

[64] J Bouix, J Krupinski, R Rzepecki, B Nowosad, I Skrzyzala, M Roborzynski, W Fudalewicz-Niemczyk, M Skalska, A Malczewski, and L Gruner. Genetic resistance to gastrointestinal nematode parasites in polish long-wool sheep. *International Journal for Parasitology*, 28(11):1797–1804, November 1998. PMID: 9846618.

[65] H Bovenhuis and R J Spelman. Selective genotyping to detect quantitative trait loci for multiple traits in outbred populations. *Journal of Dairy Science*, 83(1):173–180, January 2000. PMID: 10659977.

[66] Bruno S.A.F. Brasil, Ronaldo L. Nunes, Eduardo Bastianetto, Marcela G. Drummond, Daniel C. Carvalho, Romário C. Leite, Marcelo B. Molento, and Denise A.A. Oliveira. Genetic diversity patterns of haemonchus placei and haemonchus contortus populations isolated from domestic ruminants in brazil. *International Journal for Parasitology*, 42(5):469–479, May 2012.

[67] J M Burke, R M Kaplan, J E Miller, T H Terrill, W R Getz, S Mobini, E Valencia, M J Williams, L H Williamson, and A F Vatta. Accuracy of the FAMACHA system for on-farm use by sheep and goat producers in the southeastern united states. *Veterinary parasitology*, 147(1-2):89–95, June 2007. PMID: 17482368.

[68] J.M Burke and J.E Miller. Relative resistance of dorper crossbred ewes to gastrointestinal nematode infection compared with st. croix and katahdin ewes in the southeastern united states. *Veterinary Parasitology*, 109(3-4):265–275, November 2002.

[69] J.M Burke and J.E Miller. Relative resistance to gastrointestinal nematode parasites in dorper, katahdin, and st. croix lambs under conditions encountered in the southeastern region of the united states. *Small Ruminant Research*, 54(1-2):43–51, August 2004.

[70] J.M. Burke and J.E. Miller. Evaluation of multiple low doses of copper oxide wire particles compared with levamisole for control of haemonchus contortus in lambs. *Veterinary Parasitology*, 139(1-3):145–149, June 2006.

[71] J.M. Burke, J.E. Miller, and D.K. Brauer. The effectiveness of copper oxide wire particles as an anthelmintic in pregnant ewes and safety to offspring. *Veterinary Parasitology*, 131(3-4):291–297, August 2005.

[72] J.M. Burke, J.E. Miller, D.D. Olcott, B.M. Olcott, and T.H. Terrill. Effect of copper oxide wire particles dosage and feed supplement level on haemonchus contortus infection in lambs. *Veterinary Parasitology*, 123(3-4):235–243, September 2004.

[73] J.M. Burke, D. Morrical, and J.E. Miller. Control of gastrointestinal nematodes with copper oxide wire particles in a flock of lactating polypay ewes and offspring in iowa, USA. *Veterinary Parasitology*, 146(3-4):372–375, May 2007.

[74] F. M. Burnet. *The Clonal Selection Theory of Acquired Immunity*. Vanderbilt Univ. Press, Nashville, TN, 1959.

[75] S A Bustin. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. *Journal of molecular endocrinology*, 29(1):23–39, August 2002. PMID: 12200227.

[76] J Buítkamp, P Filmether, M J Stear, and J T Epplen. Class i and class II major histocompatibility complex alleles are associated with faecal egg counts following natural, predominantly ostertagia circumcincta infection. *Parasitology research*, 82(8):693–696, 1996. PMID: 8897503.

[77] J Cabaret. Pro and cons of targeted selective treatment against digestive-tract strongyles of ruminants. *Parasite (Paris, France)*, 15(3):506–509, September 2008. PMID: 18814731.

[78] J Cabaret, M Benoit, G Laignel, and C Nicourt. Current management of farms and internal parasites by conventional and organic meat sheep french farmers and acceptance of targeted selective treatments. *Veterinary parasitology*, 164(1):21–29, September 2009. PMID: 19414221.

[79] J. Cabaret, V. Gonnord, J. Cortet, C Sauvé, J. Ballet, H. Tournadre, and M Benoit. Indicators for internal parasitic infections in organic flocks: the diarrhoea score (Disco) proposal for lambs. Odense, Denmark, 2006.

[80] Jacques Cabaret, Michel Bouilhol, and Christian Mage. Managing helminths of ruminants in organic farming. *Veterinary research*, 33(5):625–640, October 2002. PMID: 12387494.

[81] A Carta, Sara Casu, and S Salaris. Invited review: Current state of genetic improvement in dairy sheep. *Journal of dairy science*, 92(12):5814–5833, December 2009. PMID: 19923587.

[82] Juan Antonio Figueroa Castillo, Rubén Danilo Méndez Medina, José Manuel Berruecos Villalobos, Amanda Gayosso-Vázquez, Raúl Ulloa-Arvízu, Rebeca Acosta Rodríguez, Hugo Pérez Ramírez, and Rogelio A Alonso Morales. Association between major histocompatibility complex microsatellites, fecal egg count, blood packed cell volume and blood eosinophilia in pelibuey sheep infected with haemonchus contortus. *Veterinary parasitology*, 177(3-4):339–344, May 2011. PMID: 21208746.

[83] D Cernanská, M Várady, and J Corba. A survey on anthelmintic resistance in nematode parasites of sheep in the slovak republic. *Veterinary Parasitology*, 135(1):39–45, January 2006. PMID: 16213097.

[84] A. S. Cezar, G. Toscan, G. Camillo, L. A. Sangioni, H. O. Ribas, and F. S. Vogel. Multiple resistance of gastrointestinal nematodes to nine different drugs in a sheep flock in southern brazil. *Vet Parasitol*, 173(1-2):157–60, 2010.

[85] P. Chandrawathani, N. Yusoff, L.C. Wan, A. Ham, and P.J. Waller. Total anthelmintic failure to control nematode parasites of small ruminants on government breeding farms in sabah, east malaysia. *Veterinary Research Communications*, 28(6):479–489, August 2004.

[86] N Charbonnel and J Pemberton. A long-term genetic survey of an ungulate population reveals balancing selection acting on MHC through spatial and temporal fluctuations in selection. *Heredity*, 95(5):377–388, November 2005. PMID: 16118661.

[87] J Charlier, P Dorny, B Levecke, J Demeler, G von Samson-Himmelstjerna, J Höglund, and J Vercruysse. Serum pepsinogen levels to monitor gastrointestinal nematode infections in cattle revisited. *Research in veterinary science*, 90(3):451–456, June 2011. PMID: 20667568.

[88] C Chartier, I Pors, J Hubert, D Rocheteau, C Benoit, and N Bernard. Prevalence of anthelmintic resistant nematodes in sheep and goats in western france. *Small Ruminant Research*, 29(1):33–41, June 1998.

[89] Fei Chen, Zhugong Liu, Wenhui Wu, Cristina Rozo, Scott Bowdridge, Ariel Millman, Nico Van Rooijen, Jr Urban, Joseph F, Thomas A Wynn, and William C Gause. An essential role for TH2-type responses in limiting acute tissue damage during experimental helminth infection. *Nature medicine*, 18(2):260–266, February 2012. PMID: 22245779.

[90] G A Churchill and R W Doerge. Empirical threshold values for quantitative trait mapping. *Genetics*, 138(3):963–971, November 1994. PMID: 7851788.

[91] R A Clarke, A L Burn, I Lenane, R G Windon, and K J Beh. Molecular analysis and nematode resistance association of a polymorphism at the 5' end of the sheep IgE gene. *Veterinary immunology and immunopathology*, 79(1-2):15–29, May 2001. PMID: 11356247.

[92] J. B. Cole, P. M. VanRaden, J. R. O'Connell, C. P. Van Tassell, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and G. R. Wiggans. Distribution and location of genetic effects for dairy traits. *J Dairy Sci*, 92(6):2931–46, 2009.

[93] C Colombani, P Croiseau, S Fritz, F Guillaume, A Legarra, V Ducrocq, and C Robert-Granié. A comparison of partial least squares (PLS) and sparse PLS regressions in genomic selection in french dairy cattle. *Journal of dairy science*, 95(4):2120–2131, April 2012. PMID: 22459857.

[94] D W Coltman, K Wilson, J G Pilkington, M J Stear, and J M Pemberton. A microsatellite polymorphism in the gamma interferon gene is associated with resistance to gastrointestinal nematodes in a naturally-parasitized population of soay sheep. *Parasitology*, 122(Pt 5):571–582, May 2001. PMID: 11393831.

[95] R M Connan. Inhibited development in haemonchus contortus. *Parasitology*, 71(2):239–246, October 1975. PMID: 1187184.

[96] R.L. Coop and I. Kyriazakis. Nutrition–parasite interaction. *Veterinary Parasitology*, 84(3-4):187–204, August 1999.

[97] Robert L Coop and Ilias Kyriazakis. Influence of host nutrition on the development and consequences of nematode parasitism in ruminants. *Trends in Parasitology*, 17(7):325–330, July 2001.

[98] D.W. Cooper, R.A.H. Van Oorschot, L.R. Piper, and L.F. Le Jambre. No association between the ovine leucocyte antigen (OLA) system in the australian merino and susceptibility to haemonchus contortus infection. *International Journal for Parasitology*, 19(6):695–697, 1989.

[99] L. J. Corbin, S. C. Blott, J. E. Swinburne, M. Vaudin, S. C. Bishop, and J. A. Woolliams. Linkage disequilibrium and historical effective population size in the thoroughbred horse. *Animal Genetics*, 41:8–15, December 2010.

[100] M J Coyne, G Smith, and C Johnstone. A study of the mortality and fecundity of haemonchus contortus in sheep following experimental infections. *International journal for parasitology*, 21(7):847–853, November 1991. PMID: 1774120.

[101] A. M. Crawford, K. A. Paterson, K. G. Dodds, C. Diez Tascon, P. A. Williamson, M. Roberts Thomson, S. A. Bisset, A. E. Beattie, G. J. Greer, R. S. Green, R. Wheeler, R. J. Shaw, K. Knowler, and J. C. McEwan. Discovery of quantitative trait loci for re-

sistance to parasitic nematode infection in sheep: I. analysis of outcross pedigrees. *BMC Genomics*, 7:178, 2006.

[102] G. Cringoli, L. Rinaldi, V. Veneziano, L. Mezzino, J. Vercruysse, and F. Jackson. Evaluation of targeted selective treatments in sheep in italy: Effects on faecal worm egg count and milk production in four case studies. *Veterinary Parasitology*, 164(1):36–43, September 2009.

[103] H D Daetwyler, K E Kemper, J H J van der Werf, and B J Hayes. Components of the accuracy of genomic prediction in a multi-breed sheep population. *Journal of animal science*, 90(10):3375–3384, October 2012. PMID: 23038744.

[104] Coralie Danchin-burge. Bilan de variabilité génétique de 9 races de petits ruminants laitiers et à toison. Technical Report 001172004, Institut de l'élevage, Paris, 2011.

[105] A Darvasi. Experimental strategies for the genetic dissection of complex traits in animal models. *Nature genetics*, 18(1):19–24, January 1998. PMID: 9425894.

[106] A. Darvasi, A. Weinreb, V. Minke, J. I. Weller, and M. Soller. Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. *Genetics*, 134(3):943–51, 1993.

[107] G. Davies, S. Genini, S. C. Bishop, and E. Giuffra. An assessment of opportunities to dissect host genetic variation in resistance to infectious diseases in livestock. *animal*, 3(03):415, November 2008.

[108] G. Davies, M. J. Stear, M. Benothman, O. Abuagob, A. Kerr, S. Mitchell, and S. C. Bishop. Quantitative trait loci associated with parasitic infection in scottish blackface sheep. *Heredity*, 96(3):252–8, 2006.

[109] G. Davies, M. J. Stear, and S.C. Bishop. Genetic relationships between indicator traits and nematode parasite infection levels in 6-month-old lambs. *Animal*, 80:143–150, 2005.

[110] G.-K. Dayo, S. Thevenon, D. Berthier, K. Moazami-Goudarzi, C. Denis, G. Cuny, A. Eggen, and M. Gautier. Detection of selection signatures within candidate regions underlying trypanotolerance in outbred cattle populations. *Molecular Ecology*, 18(8):1801–1813, April 2009.

[111] Dirk-Jan de Koning and Chris S Haley. Genetical genomics in humans and model organisms. *Trends in genetics: TIG*, 21(7):377–381, July 2005. PMID: 15908034.

[112] C. De La Chevrotière, C. R. Moreno, P. Jacquiet, and N. Mandonnet. La sélection génétique pour la maîtrise des strongyloses gastro-intestinales des petits ruminants. *INRA Prod. Anim.*, 24(3):221–234, 2011.

[113] Gustavo de los Campos, Hugo Naya, Daniel Gianola, José Crossa, Andrés Legarra, Eduardo Manfredi, Kent Weigel, and José Miguel Cotes. Predicting quantitative traits with regression models for dense molecular markers and pedigree. *Genetics*, 182(1):375–385, May 2009. PMID: 19293140.

[114] A. P. de Roos, B. J. Hayes, R. J. Spelman, and M. E. Goddard. Linkage disequilibrium and persistence of phase in holstein-friesian, jersey and angus cattle. *Genetics*, 179(3):1503–12, 2008.

[115] M J de Veer, J M Kemp, and E N T Meeusen. The innate host defence against nematode parasites. *Parasite immunology*, 29(1):1–9, January 2007. PMID: 17187650.

[116] Jelle de Wit, Yuri Souwer, Astrid J van Beelen, Rosa de Groot, Femke J M Muller, Hanny Klaasse Bos, Tineke Jorritsma, Martien L Kapsenberg, Esther C de Jong, and S Marieke van Ham. CD5 costimulation induces stable th17 development by promoting IL-23R expression and sustained STAT3 activation. *Blood*, 118(23):6107–6114, December 2011. PMID: 21926348.

[117] Ellen Decaestecker, Sabrina Gaba, Joost A M Raeymaekers, Robby Stoks, Liesbeth Van Kerckhoven, Dieter Ebert, and Luc De Meester. Host-parasite 'Red queen' dynamics archived in pond sediment. *Nature*, 450(7171):870–873, December 2007. PMID: 18004303.

[118] J C M Dekkers. Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. *Journal of Animal Science*, 82 E-Suppl:E313–328, 2004. PMID: 15471812.

[119] E Dervishi, J Uriarte, J Valderrábano, and J H Calvo. Structural and functional characterisation of the ovine interferon gamma (IFNG) gene: its role in nematode resistance in rasa aragonesa ewes. *Veterinary Immunology and Immunopathology*, 141(1-2):100–108, May 2011. PMID: 21419496.

[120] Johann C Detilleux. Effectiveness analysis of resistance and tolerance to infection. *Genetics, selection, evolution: GSE*, 43:9, 2011. PMID: 21362170.

[121] B. Devlin and N. Risch. A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics*, 29(2):311–22, 1995.

[122] C Diez-Tascón, O M Keane, T Wilson, A Zadissa, D L Hyndman, D B Baird, J C McEwan, and A M Crawford. Microarray analysis of selection lines from outbred populations to identify genes involved with nematode parasite resistance in sheep. *Physiological genomics*, 21(1):59–69, March 2005. PMID: 15623564.

[123] R J Dobson, E H Barnes, K L Tyrrell, B C Hosking, J W A Larsen, R B Besier, S Love, P F Rolfe, and J N Bailey. A multi-species model to assess the effect of refugia on worm control and anthelmintic resistance in sheep grazing systems. *Australian Veterinary Journal*, 89(6):200–208, June 2011. PMID: 21595639.

[124] K.G. Dodds, J.C. McEwan, and G.H. Davis. Integration of molecular and quantitative information in sheep and goat industry breeding programmes. *Small Ruminant Research*, 70(1):32–41, June 2007.

[125] A. B. Doeschl-Wilson. The role of mathematical models of host–pathogen interactions for livestock health and production – a review. *animal*, 5(06):895–910, January 2011.

[126] Andrea Beate Doeschl-Wilson, Dimitrios Vagenas, Ilias Kyriazakis, and Stephen Christopher Bishop. Exploring the assumptions underlying genetic variation in host nematode resistance (Open access publication). *Genetics, selection, evolution: GSE*, 40(3):241–264, June 2008. PMID: 18400148.

[127] S. Dominik. Quantitative trait loci for internal nematode resistance in sheep: a review. *Genet Sel Evol*, 37 Suppl 1:S83–96, 2005.

[128] S. Dominik, P. W. Hunt, J. McNally, A. Murrell, A. Hall, and I. W. Purvis. Detection of quantitative trait loci for internal parasite resistance in sheep. i. linkage analysis in a romney x merino sheep backcross population. *Parasitology*, 137(8):1275–82, 2010.

[129] P. Dorny and J. Vercruysse. Evaluation of a micro method for the routine determination of serum pepsinogen in cattle. *Res Vet Sci*, 65(3):259–62, 1998.

[130] P G Douch, R S Green, C A Morris, J C McEewan, and R G Windon. Phenotypic markers for selection of nematode-resistant sheep. *International journal for parasitology*, 26(8-9):899–911, September 1996. PMID: 8923138.

[131] E K Doyle, L P Kahn, and S J McClure. Rumen function and digestion of merino sheep divergently selected for genetic difference in resistance to haemonchus contortus. *Veterinary parasitology*, 179(1-3):130–136, June 2011. PMID: 21376471.

[132] E K Doyle, L P Kahn, S J McClure, and J M Lea. Voluntary feed intake and diet selection of merino sheep divergently selected for genetic difference in resistance to haemonchus contortus. *Veterinary parasitology*, 177(3-4):316–323, May 2011. PMID: 21330058.

[133] Tom Druet, Sébastien Fritz, Mekki Boussaha, Slim Ben-Jemaa, François Guillaume, David Derbala, Diana Zelenika, Doris Lechner, Céline Charon, Didier Boichard, Ivo G Gut, André Eggen, and Mathieu Gautier. Fine mapping of quantitative trait loci affecting female fertility in dairy cattle on BTA03 using a dense single-nucleotide polymorphism map. *Genetics*, 178(4):2227–2235, April 2008. PMID: 18430945.

[134] S I Duchemin, C Colombani, A Legarra, G Baloche, H Larroque, J-M Astruc, F Barillet, C Robert-Granié, and E Manfredi. Genomic selection in the french lacaune dairy sheep breed. *Journal of dairy science*, 95(5):2723–2733, May 2012. PMID: 22541502.

[135] V S R Dukkipati, H T Blair, D J Garrick, and A Murray. 'Ovar-Mhc' - ovine major histocompatibility complex: structure and gene polymorphisms. *Genetics and molecular research: GMR*, 5(4):581–608, 2006. PMID: 17183471.

[136] J L Dunphy, A Balic, G J Barcham, A J Horvath, A D Nash, and E N Meeusen. Isolation and characterization of a novel inducible mammalian galectin. *The Journal of biological chemistry*, 275(41):32106–32113, October 2000. PMID: 10851241.

[137] S.J Eady, R.R Woolaston, and I.A Barger. Comparison of genetic and nongenetic strategies for control of gastrointestinal nematodes of sheep. *Livestock Production Science*, 81(1):11–23, May 2003.

[138] F Echevarria, M F Borba, A C Pinheiro, P J Waller, and J W Hansen. The prevalence of anthelmintic resistance in nematode parasites of sheep in southern latin america: Brazil. *Veterinary Parasitology*, 62(3-4):199–206, April 1996. PMID: 8686165.

[139] C Eddi, J Caracostantogolo, M Peña, J Schapiro, L Marangunich, P J Waller, and J W Hansen. The prevalence of anthelmintic resistance in nematode parasites of sheep in southern latin america: Argentina. *Veterinary Parasitology*, 62(3-4):189–197, April 1996. PMID: 8686164.

[140] Melinda J Ekkens, Zhugong Liu, Qian Liu, Jeannette Whitmire, Shiyun Xiao, Anthony Foster, John Pesce, Jansie VanNoy, Arlene H Sharpe, Joseph F Urban, and William C Gause. The role of OX40 ligand interactions in the development of the th2 response to the gastrointestinal nematode parasite heligmosomoides polygyrus. *Journal of immunology (Baltimore, Md.: 1950)*, 170(1):384–393, January 2003. PMID: 12496423.

[141] Jean-Michel Elsen, Brigitte Mangin, Bruno Goffinet, Didier Boichard, and Pascale Le Roy. Alternative models for QTL detection in livestock. i. general introduction. *Genetics Selection Evolution*, 31:213, 1999.

[142] D L Emery, S J McClure, and B M Wagland. Production of vaccines against gastrointestinal nematodes of livestock. *Immunology and cell biology*, 71 (Pt 5):463–472, October 1993. PMID: 8270275.

[143] D.L. Emery. Vaccination against worm parasites of animals. *Veterinary Parasitology*, 64(1-2):31–45, August 1996.

[144] L. O. Emik. The effects of environmental and hereditary factors on trichostrongylid worm infestation in sheep. *J Anim Sci*, 8:73–80, 1949.

[145] C Epe, C Holst, R Koopmann, T Schnieder, M Larsen, and G von Samson-Himmelstjerna. Experiences with duddingtonia flagrans administration to parasitized small ruminants. *Veterinary parasitology*, 159(1):86–90, January 2009. PMID: 19019550.

[146] Klaus J Erb. Can helminths or helminth-derived products be used in humans to prevent or treat allergic diseases? *Trends in immunology*, 30(2):75–82, February 2009. PMID: 19138565.

[147] Bart Everts, Hermelijn H Smits, Cornelis H Hokke, and Maria Yazdanbakhsh. Helminths and dendritic cells: sensing and regulating via pattern recognition receptors, th2 and treg responses. *European journal of immunology*, 40(6):1525–1537, June 2010. PMID: 20405478.

[148] B. Ewing and P. Green. Analysis of expressed sequence tags indicates 35,000 human genes. *Nat Genet*, 25(2):232–4, 2000.

[149] M Eysker. Some aspects of inhibited development of trichostrongylids in ruminants. *Veterinary Parasitology*, 72(3-4):265–272; discussion 272–283, November 1997. PMID: 9460202.

[150] M. Eysker, N. Bakker, Y.A. van der Hall, I. van Hecke, F.N.J. Kooyman, D. van der Linden, C. Schrama, and H.W. Ploeger. The impact of daily duddingtonia flagrans application to lactating ewes on gastrointestinal nematodes infections in their lambs in the netherlands. *Veterinary Parasitology*, 141(1-2):91–100, October 2006.

[151] D. S Falconer and T. F. C Mackay. *Introduction to quantitative genetics*. Longman, Essex, England, 1996.

[152] F. Farnir, W. Coppieters, J. J. Arranz, P. Berzi, N. Cambisano, B. Grisart, L. Karim, F. Marcq, L. Moreau, M. Mni, C. Nezer, P. Simon, P. Vanmanshoven, D. Wagenaar, and M. Georges. Extensive genome-wide linkage disequilibrium in cattle. *Genome Res*, 10(2):220–7, 2000.

[153] Frédéric Farnir, Bernard Grisart, Wouter Coppieters, Juliette Riquet, Paulette Berzi, Nadine Cambisano, Latifa Karim, Myriam Mni, Sirja Moisio, Patricia Simon, Danny Wagenaar, Johanna Vilkki, and Michel Georges. Simultaneous mining of linkage and linkage disequilibrium to fine map quantitative trait loci in outbred half-sib pedigrees: revisiting the location of a quantitative trait locus with major effect on milk production on bovine chromosome 14. *Genetics*, 161(1):275–287, May 2002. PMID: 12019241.

[154] Z Farooq, Z Iqbal, S Mushtaq, G Muhammad, M Iqbal, and M Arshad. Ethnoveterinary practices for the treatment of parasitic diseases in livestock in cholistan desert (Pakistan). *Journal of Ethnopharmacology*, 118(2):213–219, July 2008. PMID: 18524514.

[155] J. Felsenstein. The effect of linkage on directional selection. *Genetics*, 52(2):349–63, 1965.

[156] R. L. Fernando and M. Grossman. Marker assisted selection using best linear unbiased prediction. *Genet Sel Evol*, 21:467–477, 1989.

[157] R L Fernando, D Nettleton, B R Southey, J C M Dekkers, M F Rothschild, and M Soller. Controlling the proportion of false positives in multiple dependent tests. *Genetics*, 166(1):611–619, January 2004. PMID: 15020448.

[158] M W Fleming. Size of inoculum dose regulates in part worm burdens, fecundity, and lengths in ovine haemonchus contortus infections. *The Journal of parasitology*, 74(6):975–978, December 1988. PMID: 3193336.

[159] S. A. Flint-Garcia, J. M. Thornsberry, and E. S. th Buckler. Structure of linkage disequilibrium in plants. *Annu Rev Plant Biol*, 54:357–74, 2003.

[160] S P Fodor, R P Rava, X C Huang, A C Pease, C P Holmes, and C L Adams. Multiplexed biochemical assays with biological chips. *Nature*, 364(6437):555–556, August 1993. PMID: 7687751.

[161] S P Fodor, J L Read, M C Pirrung, L Stryer, A T Lu, and D Solas. Light-directed, spatially addressable parallel chemical synthesis. *Science (New York, N.Y.)*, 251(4995):767–773, February 1991. PMID: 1990438.

[162] A B Forbes, M Warren, M Upjohn, B Jackson, J Jones, J Charlier, and M T Fox. Associations between blood gastrin, ghrelin, leptin, pepsinogen and ostertagia ostertagi antibody concentrations and voluntary feed intake in calves exposed to a trickle infection with o. ostertagi. *Veterinary parasitology*, 162(3-4):295–305, June 2009. PMID: 19342178.

[163] R H J Forrest, H Zhou, Q Fang, A Smyth, C M Frampton, and J G H Hickford. No evidence for a universal association between variation in faecal egg count for a mixed field-challenge of gastrointestinal parasites and the presence of the ovar-DQA1 null haplotype in sheep. *Veterinary immunology and immunopathology*, 135(3-4):303–305, June 2010. PMID: 20097431.

[164] M. T. Fox. Pathophysiology of infection with gastrointestinal nematodes in domestic ruminants: recent developments. *Vet Parasitol*, 72(3-4):285–97; discussion 297–308, 1997.

[165] Anne T French, Julie A Bethune, Pamela A Knight, Tom N McNeilly, Sean Wattegedera, Susan Rhind, Hugh R P Miller, and Alan D Pemberton. The expression of intelectin in sheep goblet cells and upregulation by interleukin-4. *Veterinary immunology and immunopathology*, 120(1-2):41–46, November 2007. PMID: 17727963.

[166] Anne T French, Pamela A Knight, W David Smith, Jeremy K Brown, Nicola M Craig, Judith A Pate, Hugh R P Miller, and Alan D Pemberton. Up-regulation of intelectin in

sheep after infection with teladorsagia circumcincta. *International journal for parasitology*, 38(3-4):467–475, March 2008. PMID: 17983620.

[167] S Gaba, J Cabaret, C Chylinski, C Sauvé, J Cortet, and A Silvestre. Can efficient management of sheep gastro-intestinal nematodes be based on random treatment? *Veterinary parasitology*, June 2012. PMID: 22770705.

[168] S Gaba, J Cabaret, C Sauvé, J Cortet, and A Silvestre. Experimental and modeling approaches to evaluate different aspects of the efficacy of targeted selective treatment of anthelmintics against sheep parasite nematodes. *Veterinary Parasitology*, 171(3-4):254–262, August 2010. PMID: 20430530.

[169] S. Gaba, V. Ginot, and J. Cabaret. Modelling macroparasite aggregation using a nematode-sheep system: the weibull distribution as an alternative to the negative binomial distribution? *Parasitology*, 131(03):393, April 2005.

[170] S Gabrie, I K Phiri, P Dorny, and J Vercruysse. A survey on anthelmintic resistance in nematode parasites of sheep in lusaka, zambia. *The Onderstepoort Journal of Veterinary Research*, 68(4):271–274, December 2001. PMID: 12026061.

[171] Stephen J Galli and Mindy Tsai. Mast cells in allergy and infection: versatile effector and regulatory cells in innate and adaptive immunity. *European journal of immunology*, 40(7):1843–1851, July 2010. PMID: 20583030.

[172] E Gallidis, E Papadopoulos, S Ptochos, and G Arsenos. The use of targeted selective treatments against gastrointestinal nematodes in milking sheep and goats in greece based on parasitological and performance criteria. *Veterinary parasitology*, 164(1):53–58, September 2009. PMID: 19409707.

[173] M. Gauly, M. Kraus, L. Vervelde, M.A.W. van Leeuwen, and G. Erhardt. Estimating genetic differences in natural resistance in rhön and merinoland sheep following experimental haemonchus contortus infection. *Veterinary Parasitology*, 106(1):55–67, May 2002.

[174] Mathieu Gautier, Thomas Faraut, Katayoun Moazami-Goudarzi, Vincent Navratil, Mario Foglio, Cécile Grohs, Anne Boland, Jean-Guillaume Garnier, Didier Boichard, G Mark Lathrop, Ivo G Gut, and André Eggen. Genetic and haplotypic structure in 14 european and african cattle breeds. *Genetics*, 177(2):1059–1070, October 2007. PMID: 17720924.

[175] C Genchi. [In memory of ivo de carneri (1927-1993)]. *Parassitologia*, 36(3):229–242, December 1994. PMID: 7637992.

[176] Michel Georges. Mapping, fine mapping, and molecular dissection of quantitative trait loci in domestic animals. *Annual Review of Genomics and Human Genetics*, 8:131–162, 2007. PMID: 17477823.

[177] Paul R Giacomin, David L Gordon, Marina Botto, Mohamed R Daha, Sam D Sanderson, Stephen M Taylor, and Lindsay A Dent. The role of complement in innate, adaptive and eosinophil-dependent immunity to the nematode nippostrongylus brasiliensis. *Molecular immunology*, 45(2):446–455, January 2008. PMID: 17675237.

[178] Daniel Gianola, Gustavo de los Campos, William G Hill, Eduardo Manfredi, and Rohan Fernando. Additive genetic variability and the bayesian alphabet. *Genetics*, 183(1):347–363, September 2009. PMID: 19620397.

[179] Daniel Gianola, Miguel Perez-Enciso, and Miguel A Toro. On marker-assisted prediction of genetic value: beyond the ridge. *Genetics*, 163(1):347–365, January 2003. PMID: 12586721.

[180] R A Gibbs, J F Taylor, Curtis P Van Tassell, William Barendse, Kellye A Eversole, Clare A Gill, Ronnie D Green, Debora L Hamernik, Steven M Kappes, Sigbjørn Lien, Lakshmi K Matukumalli, John C McEwan, Lynne V Nazareth, Robert D Schnabel, George M Weinstock, David A Wheeler, Paolo Ajmone-Marsan, Paul J Boettcher, Alexandre R Caetano, Jose Fernando Garcia, Olivier Hanotte, Paola Mariani, Loren C Skow, Tad S Sonstegard, John L Williams, Boubacar Diallo, Lemecha Hailemariam, Mario L Martinez, Chris A Morris, Luiz O C Silva, Richard J Spelman, Woudyalew Mulatu, Keyan Zhao, Colette A Abbey, Morris Agaba, Flábio R Araujo, Rowan J Bunch, James Burton, Chiara Gorni, Hanotte Olivier, Blair E Harrison, Bill Luff, Marco A Machado, Joel Mwakaya, Graham Plastow, Warren Sim, Timothy Smith, Merle B Thomas, Alessio Valentini, Paul Williams, James Womack, John A Woolliams, Yue Liu, Xiang Qin, Kim C Worley, Chuan Gao, Huaiyang Jiang, Stephen S Moore, Yanru Ren, Xing-Zhi Song, Carlos D Bustamante, Ryan D Hernandez, Donna M Muzny, Shobha Patil, Anthony San Lucas, Qing Fu, Matthew P Kent, Richard Vega, Aruna Matukumalli, Sean McWilliam, Gert Sclep, Katarzyna Bryc, Jungwoo Choi, Hong Gao, John J Grefenstette, Brenda Murdoch, Alessandra Stella, Rafael

Villa-Angulo, Mark Wright, Jan Aerts, Oliver Jann, Riccardo Negrini, Mike E Goddard, Ben J Hayes, Daniel G Bradley, Marcos Barbosa da Silva, Lilian P L Lau, George E Liu, David J Lynn, Francesca Panzitta, and Ken G Dodds. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science (New York, N.Y.)*, 324(5926):528–532, April 2009. PMID: 19390050.

[181] M G Gicheha, I S Kosgey, B O Bebe, and A K Kahi. Evaluation of the efficiency of alternative two-tier nucleus breeding systems designed to improve meat sheep in kenya. *Journal of animal breeding and genetics = Zeitschrift für Tierzüchtung und Züchtungsbiologie*, 123(4):247–257, August 2006. PMID: 16882091.

[182] Hélène Gilbert and Pascale Le Roy. Methods for the detection of multiple linked QTL applied to a mixture of full and half sib families. *Genetics, selection, evolution: GSE*, 39(2):139–158, April 2007. PMID: 17306198.

[183] B S Gill. Anthelmintic resistance in india. *Veterinary Parasitology*, 63(1-2):173–176, May 1996. PMID: 8792590.

[184] H S Gill, K Altmann, M L Cross, and A J Husband. Induction of t helper 1- and t helper 2-type immune responses during haemonchus contortus infection in sheep. *Immunology*, 99(3):458–463, March 2000. PMID: 10712677.

[185] John Stuart Gilleard. Understanding anthelmintic resistance: the need for genomics and genetics. *International journal for parasitology*, 36(12):1227–1239, October 2006. PMID: 16889782.

[186] J B Githiori, S Athanasiadou, and S M Thamsborg. Use of plants in novel approaches for control of gastrointestinal helminths in livestock with emphasis on small ruminants. *Veterinary Parasitology*, 139(4):308–320, July 2006.

[187] M.E. Goddard and B.J. Hayes. Genomic selection. *Journal of Animal Breeding and Genetics*, 124(6):323–330, December 2007.

[188] Michael E Goddard and Ben J Hayes. Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nature reviews. Genetics*, 10(6):381–391, June 2009. PMID: 19448663.

[189] B. Goffinet and B. Mangin. Comparing methods to detect more than one QTL on a chromosome. *TAG Theoretical and Applied Genetics*, 96(5):628–633, April 1998.

[190] D E Goldgar. Multipoint analysis of human quantitative genetic variation. *American journal of human genetics*, 47(6):957–967, December 1990. PMID: 2239972.

[191] Jorge F González, Alvaro Hernández, Els N T Meeusen, Francisco Rodríguez, José M Molina, José R Jaber, Herman W Raadsma, and David Piedrafita. Fecundity in adult haemonchus contortus parasites is correlated with abomasal tissue eosinophils and $\gamma\delta$ t cells in resistant canaria hair breed sheep. *Veterinary Parasitology*, 178(3-4):286–292, June 2011. PMID: 21277683.

[192] Jorge F González, Alvaro Hernández, José M Molina, Antonio Fernández, Herman W Raadsma, Els N T Meeusen, and David Piedrafita. Comparative experimental haemonchus contortus infection of two sheep breeds native to the canary islands. *Veterinary parasitology*, 153(3-4):374–378, May 2008. PMID: 18394807.

[193] B. Good, J. P. Hanrahan, B. A. Crowley, and G. Mulcahy. Texel sheep are more resistant to natural nematode challenge than suffolk sheep based on faecal egg count and nematode burden. *Vet Parasitol*, 136(3-4):317–27, 2006.

[194] B. Goossens, S. Osaer, M. Ndao, J. Van Winghem, and S. Geerts. The susceptibility of djallonké and djallonké-sahelian crossbred sheep to trypanosoma congolense and helminth infection under different diet levels. *Veterinary Parasitology*, 85(1):25–41, August 1999.

[195] J.T. Gradé, J R Tabuti, and P Van Damme. Ethnoveterinary knowledge in pastoral karamoja, uganda. *Journal of Ethnopharmacology*, 122(2):273–293, March 2009.

[196] Melissa M Gray, Julie M Granka, Carlos D Bustamante, Nathan B Sutter, Adam R Boyko, Lan Zhu, Elaine A Ostrander, and Robert K Wayne. Linkage disequilibrium and demographic history of wild and domestic canids. *Genetics*, 181(4):1493–1505, April 2009. PMID: 19189949.

[197] P. W. Gregory. The possibility of establishing within breeds lines of sheep that are genetically resistant to stomach worms. *Am. Soc. Anim. Prod.*, pages 316–324, 1937.

[198] L. Gruner, G. Aumont, T. Getachew, J. C. Brunel, C. Pery, Y. Cognie, and Y. Guerin. Experimental infection of black belly and INRA 401 straight and crossbred sheep with trichostrongyle nematode parasites. *Vet Parasitol*, 116(3):239–49, 2003.

[199] L Gruner, J Bouix, and J C Brunel. High genetic correlation between resistance to haemonchus contortus and to trichostrongylus colubriformis in INRA 401 sheep. *Veterinary Parasitology*, 119(1):51–58, January 2004. PMID: 15036576.

[200] L Gruner, J Cortet, C Sauvé, C Limouzin, and J.C Brunel. Evolution of nematode community in grazing sheep selected for resistance and susceptibility to teladorsagia circumcincta and trichostrongylus colubriformis: a 4-year experiment. *Veterinary Parasitology*, 109(3-4):277–291, November 2002.

[201] Lucas Gruner, Jacques Bouix, Jacqueline Vu Tien Khang, Nathalie Mandonnet, Francis Eychenne, Jacques Cortet, Christine Sauvé, and Claude Limouzin. A short-term divergent selection for resistance to teladorsagia circumcincta in romanov sheep using natural or artificial challenge. *Genetics, Selection, Evolution: GSE*, 36(2):217–242, April 2004. PMID: 15040900.

[202] François Guillaume, Sébastien Fritz, Didier Boichard, and Tom Druet. Estimation by simulation of the efficiency of the french marker-assisted selection program in dairy cattle. *Genetics, selection, evolution: GSE*, 40(1):91–102, February 2008. PMID: 18096117.

[203] B. Gutierrez-Gil, J. Perez, L. Alvarez, M. Martinez-Valladares, L. F. de la Fuente, Y. Bayon, A. Meana, F. San Primitivo, F. A. Rojo-Vazquez, and J. J. Arranz. Quantitative trait loci for resistance to trichostrongylid infection in spanish churra sheep. *Genet Sel Evol*, 41:46, 2009.

[204] B. Gutiérrez-Gil, J. Pérez, L. F. de la Fuente, A. Meana, M. Martínez-Valladares, F. San Primitivo, F. A. Rojo-Vázquez, and J. J. Arranz. Genetic parameters for resistance to trichostrongylid infection in dairy sheep. *animal*, 4(04):505, November 2009.

[205] D Habier, R L Fernando, K Kizilkaya, and D J Garrick. Extension of the bayesian alphabet for genomic selection. *BMC bioinformatics*, 12:186, 2011. PMID: 21605355.

[206] Aynalem Haile, S Tembely, D.O Anindo, E Mukasa-Mugerwa, J.E.O Rege, Alemu Yami, and R.L Baker. Effects of breed and dietary protein supplementation on the responses

to gastrointestinal nematode infections in ethiopian sheep. *Small Ruminant Research*, 44(3):247–261, June 2002.

[207] C S Haley and S A Knott. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity*, 69(4):315–324, 1992. PMID: 16718932.

[208] C S Haley, S A Knott, and J M Elsen. Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics*, 136(3):1195–1207, March 1994. PMID: 8005424.

[209] N Harmegnies, F Farnir, F Davin, N Buys, M Georges, and W Coppieters. Measuring the extent of linkage disequilibrium in commercial pig populations. *Animal Genetics*, 37(3):225–231, June 2006. PMID: 16734681.

[210] G B L Harrison, H D Pulford, E E Doolin, A Pernthaner, C B Shoemaker, and W R Hein. Antibodies to surface epitopes of the carbohydrate larval antigen CarLA are associated with passive protection in strongylid nematode challenge infections. *Parasite immunology*, 30(11-12):577–584, December 2008. PMID: 19067838.

[211] G B L Harrison, H D Pulford, W R Hein, T K Barber, R J Shaw, M McNeill, St J Wakefield, and C B Shoemaker. Immune rejection of trichostrongylus colubriformis in sheep; a possible role for intestinal mucus antibody against an l3-specific surface antigen. *Parasite immunology*, 25(1):45–53, January 2003. PMID: 12753437.

[212] G B L Harrison, H D Pulford, W R Hein, W B Severn, and C B Shoemaker. Characterization of a 35-kDa carbohydrate larval antigen (CarLA) from trichostrongylus colubriformis; a potential target for host immunity. *Parasite immunology*, 25(2):79–86, February 2003. PMID: 12791103.

[213] D. L. Hartl and A. G. Clark. *Principles of population genetics*. Sinauer Associates, Inc., Sunderland, MA, 2007.

[214] Musa Hassan, Barbara Good, James P Hanrahan, Deirdre Campion, Gearoid Sayers, Grace Mulcahy, and Torres Sweeney. The dynamic influence of the DRB1*1101 allele on the resistance of sheep to experimental teladorsagia circumcincta infection. *Veterinary research*, 42(1):46, February 2011. PMID: 21385412.

[215] Musa Hassan, James P Hanrahan, Barbara Good, Grace Mulcahy, and Torres Sweeney. A differential interplay between the expression of Th1/Th2/Treg related cytokine genes in teladorsagia circumcincta infected DRB1*1101 carrier lambs. *Veterinary research*, 42(1):45, February 2011. PMID: 21385411.

[216] B. Hayes and M. E. Goddard. The distribution of the effects of genes affecting quantitative traits in livestock. *Genet Sel Evol*, 33(3):209–29, 2001.

[217] B J Hayes, P J Bowman, A J Chamberlain, and M E Goddard. Invited review: Genomic selection in dairy cattle: progress and challenges. *Journal of dairy science*, 92(2):433–443, February 2009. PMID: 19164653.

[218] B J Hayes, A J Chamberlain, H McPartlan, I Macleod, L Sethuraman, and M E Goddard. Accuracy of marker-assisted selection with single markers and marker haplotypes in cattle. *Genetical research*, 89(4):215–220, August 2007. PMID: 18208627.

[219] P. W. Hedrick. Gametic disequilibrium measures: proceed with caution. *Genetics*, 117(2):331–41, 1987.

[220] C A Heid, J Stevens, K J Livak, and P M Williams. Real time quantitative PCR. *Genome research*, 6(10):986–994, October 1996. PMID: 8908518.

[221] E M Heifetz, J E Fulton, N O'Sullivan, H Zhao, J C M Dekkers, and M Soller. Extent and consistency across generations of linkage disequilibrium in commercial layer chicken breeding populations. *Genetics*, 171(3):1173–1181, November 2005. PMID: 16118198.

[222] W R Hein and G B L Harrison. Vaccines against veterinary helminths. *Veterinary parasitology*, 132(3-4):217–222, September 2005. PMID: 16095841.

[223] W R Hein, A Pernthaner, D Piedrafita, and E N Meeusen. Immune mechanisms of resistance to gastrointestinal nematode infections in sheep. *Parasite Immunology*, 32(8):541–548, August 2010. PMID: 20626809.

[224] H Herlich. A digestion method for post-mortem recovery of nematodes from ruminants. *Proceedings of the Helminthological Society of Washington*, 23:102–103, 1956.

[225] H Hertzberg and C Bauer. [Anthelmintic resistance in gastrointestinal strongylidae in sheep and goats: new data on prevalence, epidemiology, preventive measures and alter-

natives to anthelmintic drugs]. *Berliner Und Münchener Tierärztliche Wochenschrift*, 113(4):122–128, April 2000. PMID: 10816910.

[226] James P. Hewitson, John R. Grainger, and Rick M. Maizels. Helminth immunoregulation: The role of parasite secreted proteins in modulating host immunity. *Molecular and Biochemical Parasitology*, 167(1):1–11, September 2009.

[227] Johan Höglund, Katarina Gustafsson, Britt-Louise Ljungström, Annie Engström, Alison Donnan, and Philip Skuce. Anthelmintic resistance in swedish sheep flocks based on a comparison of the results from the faecal egg count reduction test and resistant allele frequencies of the beta-tubulin gene. *Veterinary Parasitology*, 161(1-2):60–68, April 2009. PMID: 19155135.

[228] J G H Hickford, R H J Forrest, H Zhou, Q Fang, and C M Frampton. Association between variation in faecal egg count for a mixed field-challenge of nematode parasites and ovine MHC-DQA2 polymorphism. *Veterinary immunology and immunopathology*, 144(3-4):312–320, December 2011. PMID: 21945177.

[229] A Hielscher, H Brandt, G Erhardt, and M Gauly. Heterosis analysis of haemonchus contortus resistance and production traits in rhoen sheep, merino land sheep and crossbred lambs. *Veterinary parasitology*, 141(3-4):279–284, November 2006. PMID: 16828228.

[230] R Higuchi, G Dollinger, P S Walsh, and R Griffith. Simultaneous amplification and detection of specific DNA sequences. *Bio/technology (Nature Publishing Company)*, 10(4):413–417, April 1992. PMID: 1368485.

[231] R Higuchi, C Fockler, G Dollinger, and R Watson. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Bio/technology (Nature Publishing Company)*, 11(9):1026–1030, September 1993. PMID: 7764001.

[232] W. G. Hill. Estimation of linkage disequilibrium in randomly mating populations. *Heredity*, 33(2):229–39, 1974.

[233] W. G. Hill and A Robertson. Linkage disequilibrium in finite populations. *Theoretical and Applied Genetics*, 38(6):226–231, June 1968.

[234] Joel N Hirschhorn and Mark J Daly. Genome-wide association studies for common diseases and complex traits. *Nature reviews. Genetics*, 6(2):95–108, February 2005. PMID: 15716906.

[235] M.A. Hohenhaus and P.M. Outteridge. The immunogenetics of resistance toTrichostrongylus colubriformis and haemonchus contortus parasites in sheep. *British Veterinary Journal*, 151(2):119–140, March 1995.

[236] H Hoste and J F J Torres-Acosta. Non chemical control of helminths in ruminants: adapting solutions for changing worms in a changing world. *Veterinary Parasitology*, 180(1-2):144–154, August 2011. PMID: 21705144.

[237] Hervé Hoste, Frank Jackson, Spiridoula Athanasiadou, Stig. M. Thamsborg, and Simone O. Hoskin. The effects of tannin-rich plants on parasitic nematodes in ruminants. *Trends in Parasitology*, 22(6):253–261, June 2006.

[238] David Houle, Diddahally R Govindaraju, and Stig Omholt. Phenomics: the next challenge. *Nature reviews. Genetics*, 11(12):855–866, December 2010. PMID: 21085204.

[239] S. B. Howell, J. M. Burke, J. E. Miller, T. H. Terrill, E. Valencia, M. J. Williams, L. H. Williamson, A. M. Zajac, and R. M. Kaplan. Prevalence of anthelmintic resistance on sheep and goat farms in the southeastern united states. *J Am Vet Med Assoc*, 233(12):1913–9, 2008.

[240] Zhi-Liang Hu and James M Reecy. Animal QTLdb: beyond a repository. a public platform for QTL comparisons and integration with diverse types of structural genomic information. *Mammalian genome: official journal of the International Mammalian Genome Society*, 18(1):1–4, January 2007. PMID: 17245610.

[241] R R Hudson, M Kreitman, and M Aguadé. A test of neutral molecular evolution based on nucleotide data. *Genetics*, 116(1):153–159, May 1987. PMID: 3110004.

[242] D J Hulme, F W Nicholas, R G Windon, S C Brown, and K J Beh. The MHC class II region and resistance to an intestinal parasite in sheep. *Journal of animal breeding and genetics = Zeitschrift für Tierzüchtung und Züchtungsbiologie*, 110(1-6):459–472, January 1993. PMID: 21395744.

[243] M X Ierna, H E Scales, H Schwarz, C Bunce, A McIlgorm, P Garside, and C E Lawrence. OX40 interactions in gastrointestinal nematode infection. *Immunology*, 117(1):108–116, January 2006. PMID: 16423046.

[244] Aaron Ingham, Moira Menzies, Peter Hunt, Antonio Reverter, Ross Windon, and Nicholas Andronicos. Divergent ghrelin expression patterns in sheep genetically resistant or susceptible to gastrointestinal nematodes. *Veterinary parasitology*, 181(2-4):194–202, September 2011. PMID: 21640486.

[245] Aaron Ingham, Antonio Reverter, Ross Windon, Peter Hunt, and Moira Menzies. Gastrointestinal nematode challenge induces some conserved gene expression changes in the gut mucosa of genetically resistant sheep. *International journal for parasitology*, 38(3-4):431–442, March 2008. PMID: 17826780.

[246] A Jabbar, Z Iqbal, D Kerboeuf, G Muhammad, M Khan, and M Afaq. Anthelmintic resistance: The state of play revisited. *Life Sciences*, 79(26):2413–2431, November 2006.

[247] F. Jackson, D. Bartley, Y. Bartley, and F. Kenyon. Worm control in sheep in the future. *Small Ruminant Research*, 86(1-3):40–45, October 2009.

[248] H J Jacobs, C Wiltshire, K Ashman, and E N Meeusen. Vaccination against the gastrointestinal nematode, haemonchus contortus, using a purified larval surface antigen. *Vaccine*, 17(4):362–368, January 1999. PMID: 9987175.

[249] P Jacquiet, J Cabaret, D Cheikh, and A Thiam. Experimental study of survival strategy of haemonchus contortus in sheep during the dry season in desert areas of the mauritania. *The Journal of parasitology*, 81(6):1013–1015, December 1995. PMID: 8544042.

[250] R C Jansen. Interval mapping of multiple quantitative trait loci. *Genetics*, 135(1):205–211, September 1993. PMID: 8224820.

[251] R C Jansen and J P Nap. Genetical genomics: the added value from segregation. *Trends in genetics: TIG*, 17(7):388–391, July 2001. PMID: 11418218.

[252] M. Janssen, C. Weimann, M. Gauly, and G. Erhardt. Associations between infections with haemonchus contortus and genetic markers on ovine chromosome 20. Montpellier, 2002.

[253] Marten Jäger, Claus-Eric Ott, Johannes Grünhagen, Jochen Hecht, Hanna Schell, Stefan Mundlos, Georg N Duda, Peter N Robinson, and Jasmin Lienau. Composite transcriptome

assembly of RNA-seq data in a sheep model for delayed bone healing. *BMC genomics*, 12:158, 2011. PMID: 21435219.

[254] L. B. Jorde. Linkage disequilibrium and the search for complex disease genes. *Genome Res*, 10(10):1435–44, 2000.

[255] H N Kadarmideen, N S Watson-Haigh, and N M Andronicos. Systems biology of ovine intestinal parasite resistance: disease gene modules and biomarkers. *Molecular bioSystems*, 7(1):235–246, January 2011. PMID: 21072409.

[256] R. Kaminsky, N. Gauvry, S. Schorderet Weber, T. Skripsky, J. Bouvier, A. Wenger, F. Schroeder, Y. Desaules, R. Hotz, T. Goebel, B. C. Hosking, F. Pautrat, S. Wieland-Berghausen, and P. Ducray. Identification of the amino-acetonitrile derivative monepantel (AAD 1566) as a new anthelmintic drug development candidate. *Parasitology Research*, 103(4):931–939, July 2008.

[257] Hyun Min Kang, Jae Hoon Sul, Susan K Service, Noah A Zaitlen, Sit-Yee Kong, Nelson B Freimer, Chiara Sabatti, and Eleazar Eskin. Variance component model to account for sample structure in genome-wide association studies. *Nature genetics*, 42(4):348–354, April 2010. PMID: 20208533.

[258] C H Kao. On the differences between maximum likelihood and regression interval mapping in the analysis of quantitative trait loci. *Genetics*, 156(2):855–865, October 2000. PMID: 11014831.

[259] C. H. Kao, Z. B. Zeng, and R. D. Teasdale. Multiple interval mapping for quantitative trait loci. *Genetics*, 152(3):1203–16, 1999.

[260] Ray M Kaplan. Drug resistance in nematodes of veterinary importance: a status report. *Trends in Parasitology*, 20(10):477–481, October 2004. PMID: 15363441.

[261] Ray M. Kaplan and Anand N. Vidyashankar. An inconvenient truth: Global worming and anthelmintic resistance. *Veterinary Parasitology*, November 2011.

[262] L. J. E. Karlsson and J. C. Greeff. Selection response in fecal worm egg counts in the rylington merino parasite resistant flock. *Australian Journal of Experimental Agriculture*, 46(7):809, 2006.

[263] L.J.E. Karlsson and J.C. Greeff. Genetic aspects of sheep parasitic diseases. *Veterinary Parasitology*, In press, March 2012.

[264] Fotini K Kavvoura and John P A Ioannidis. Methods for meta-analysis in genetic association studies: a review of their potential and pitfalls. *Human genetics*, 123(1):1–14, February 2008. PMID: 18026754.

[265] O M Keane, K G Dodds, A M Crawford, and J C McEwan. Identifying genes for intestinal nematode resistance using transcriptional profiling. *Developments in biologicals*, 132:205–212, 2008. PMID: 18817303.

[266] Orla M Keane, Ken G Dodds, Allan M Crawford, and John C McEwan. Transcriptional profiling of ovis aries identifies ovar-DQA1 allele frequency differences between nematode-resistant and susceptible selection lines. *Physiological genomics*, 30(3):253–261, August 2007. PMID: 17488886.

[267] Orla M Keane, Amonida Zadissa, Theresa Wilson, Dianne L Hyndman, Gordon J Greer, David B Baird, Alan F McCulloch, Allan M Crawford, and John C McEwan. Gene expression profiling of naïve sheep genetically resistant and susceptible to gastrointestinal nematodes. *BMC genomics*, 7:42, 2006. PMID: 16515715.

[268] Joanna M Kemp, Nicholas A Robinson, Els N T Meeusen, and David M Piedrafita. The relationship between the rapid rejection of haemonchus contortus larvae with cells and mediators in abomasal tissues in immune sheep. *International journal for parasitology*, 39(14):1589–1594, December 2009. PMID: 19540240.

[269] K E Kemper, R L Elwin, S C Bishop, M E Goddard, and R R Woolaston. Haemonchus contortus and trichostrongylus colubriformis did not adapt to long-term exposure to sheep that were genetically resistant or susceptible to nematode infections. *International journal for parasitology*, 39(5):607–614, April 2009. PMID: 19027020.

[270] K. E. Kemper, D. L. Emery, S. C. Bishop, H. Oddy, B. J. Hayes, S. Dominik, J. M. Henshall, and M. E. Goddard. The distribution of SNP marker effects for faecal worm egg count in sheep, and the feasibility of using these markers to predict genetic merit for resistance to worm infections. *Genet Res*, 93:203–219, 2011.

[271] K.E. Kemper, D.G. Palmer, S.M. Liu, J.C. Greeff, S.C. Bishop, and L.J.E. Karlsson. Reduction of faecal worm egg count, worm numbers and worm fecundity in sheep selected for worm resistance following artificial infection with teladorsagia circumcincta and trichostrongylus colubriformis. *Veterinary Parasitology*, 171(3-4):238–246, August 2010.

[272] F Kenyon, A W Greer, G C Coles, G Cringoli, E Papadopoulos, J Cabaret, B Berrag, M Varady, J A Van Wyk, E Thomas, J Vercruysse, and F Jackson. The role of targeted selective treatments in the development of refugia-based approaches to the control of gastrointestinal nematodes of small ruminants. *Veterinary Parasitology*, 164(1):3–11, September 2009. PMID: 19450930.

[273] F Kenyon and F Jackson. Targeted flock/herd and individual ruminant treatment approaches. *Veterinary parasitology*, 186(1-2):10–17, May 2012. PMID: 22188982.

[274] J. W. Kijas, J. A. Lenstra, B. J. Hayes, S. Boitard, L. R. Porto Neto, M.i San Cristobal, B. Servin, R. McCulloch, V. Whan, K. Gietzen, Samuel Paiva, William Barendse, Elena Ciani, Herman Raadsma, John McEwan, Brian Dalrymple, and {other members of the International Sheep Genomics Consortium}. Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biology*, 10(2):e1001258, February 2012.

[275] Yuseob Kim and Wolfgang Stephan. Detecting a local signature of genetic hitchhiking along a recombining chromosome. *Genetics*, 160(2):765–777, February 2002. PMID: 11861577.

[276] M Kimura. Evolutionary rate at the molecular level. *Nature*, 217(5129):624–626, February 1968. PMID: 5637732.

[277] Pamela A Knight, Susan E Griffith, Alan D Pemberton, Judith M Pate, Lauren Guarneri, Katherine Anderson, Richard T Talbot, Sarah Smith, David Waddington, Mark Fell, Alan L Archibald, Stewart Tg Burgess, David W Smith, Hugh Rp Miller, and Ivan W Morrison. Novel gene expression responses in the ovine abomasal mucosa to infection with the gastric nematode teladorsagia circumcincta. *Veterinary research*, 42(1):78, February 2011. PMID: 21682880.

[278] Sara A Knott. Regression-based quantitative trait loci mapping: robust, efficient and effective. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 360(1459):1435–1442, 2005. PMID: 16048786.

[279] D P Knox and W D Smith. Vaccination against gastrointestinal nematode parasites of ruminants using gut-expressed antigens. *Veterinary parasitology*, 100(1-2):21–32, September 2001. PMID: 11522403.

[280] David P. Knox, Diane L. Redmond, George F. Newlands, Philip J. Skuce, Dan Pettit, and W.David Smith. The nature and prospects for gut membrane proteins as vaccine candidates for haemonchus contortus and other ruminant trichostrongyloids. *International Journal for Parasitology*, 33(11):1129–1137, September 2003.

[281] Mr Knox. Effectiveness of copper oxide wire particles for haemonchus contortus control in sheep. *Australian Veterinary Journal*, 80(4):224–227, April 2002.

[282] M.R. Knox, J.F.J. Torres-Acosta, and A.J. Aguilar-Caballero. Exploiting the effect of dietary supplementation of small ruminants on resilience and resistance against gastrointestinal nematodes. *Veterinary Parasitology*, 139(4):385–393, July 2006.

[283] M Kojima, H Hosoda, Y Date, M Nakazato, H Matsuo, and K Kangawa. Ghrelin is a growth-hormone-releasing acylated peptide from stomach. *Nature*, 402(6762):656–660, December 1999. PMID: 10604470.

[284] W.M. Koné and K. Kamanzi Atindehou. Ethnobotanical inventory of medicinal plants used in traditional veterinary medicine in northern côte d'Ivoire (West africa). *South African Journal of Botany*, 74(1):76–84, January 2008.

[285] L Kruglyak. What is significant in whole-genome linkage disequilibrium studies? *American journal of human genetics*, 61(4):810–812, October 1997. PMID: 9382090.

[286] L Kruglyak. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nature genetics*, 22(2):139–144, June 1999. PMID: 10369254.

[287] Leonid Kruglyak. The road to genome-wide association studies. *Nature reviews. Genetics*, 9(4):314–318, April 2008. PMID: 18283274.

[288] C. Lacroux, T. H. Nguyen, O. Andreoletti, F. Prevot, C. Grisez, J. P. Bergeaud, L. Gruner, J. C. Brunel, D. Francois, P. Dorchies, and P. Jacquiet. Haemonchus contortus (Nematoda:

trichostrongylidae) infection in lambs elicits an unequivocal th2 immune response. *Vet Res*, 37(4):607–22, 2006.

[289] Steven T Laing, Al Ivens, Victoria Butler, Sai P Ravikumar, Roz Laing, Debra J Woods, and John S Gilleard. The transcriptional response of caenorhabditis elegans to ivermectin exposure identifies novel genes involved in the response to reduced food intake. *PloS One*, 7(2):e31367, 2012. PMID: 22348077.

[290] R Lande and R Thompson. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics*, 124(3):743–756, March 1990. PMID: 1968875.

[291] E. Lander and L. Kruglyak. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet*, 11(3):241–7, 1995.

[292] E. S. Lander and D. Botstein. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, 121(1):185–99, 1989.

[293] Jobst Landgrebe, Frank Bretz, and Edgar Brunner. Efficient two-sample designs for microarray experiments with biological replications. *In silico biology*, 4(4):461–470, 2004. PMID: 15506995.

[294] Denis M Larkin, Hans D Daetwyler, Alvaro G Hernandez, Chris L Wright, Lorie A Hetrick, Lisa Boucek, Sharon L Bachman, Mark R Band, Tatsiana V Akraiko, Miri Cohen-Zinder, Jyothi Thimmapuram, Iona M Macleod, Timothy T Harkins, Jennifer E McCague, Michael E Goddard, Ben J Hayes, and Harris A Lewin. Whole-genome resequencing of two elite sires for the detection of haplotypes under selection in dairy cattle. *Proceedings of the National Academy of Sciences of the United States of America*, April 2012. PMID: 22529356.

[295] J W Larsen, N Anderson, A L Vizard, G A Anderson, and H Hoste. Diarrhoea in merino ewes during winter: association with trichostrongylid larvae. *Australian veterinary journal*, 71(11):365–372, November 1994. PMID: 7726760.

[296] J W Larsen, A L Vizard, and N Anderson. Production losses in merino ewes and financial penalties caused by trichostrongylid infections during winter and spring. *Australian Veterinary Journal*, 72(2):58–63, February 1995. PMID: 7779036.

[297] M. Larsen. Biological control of helminths. *International Journal for Parasitology*, 29(1):139–146, January 1999.

[298] M Larsen. Biological control of nematode parasites in sheep. *Journal of animal science*, 84 Suppl:E133–139, April 2006. PMID: 16582084.

[299] K.E. Lawrence and R. Jackson. Risk factors for macrocyclic lactone anthelmintic resistance on new zealand sheep farms. In *Proceedings of the 11th International Symposium on Veterinary Epidemiology and Economics*, 2006.

[300] L F Le Jambre, R J Dobson, I J Lenane, and E H Barnes. Selection for anthelmintic resistance by macrocyclic lactones in haemonchus contortus. *International Journal for Parasitology*, 29(7):1101–1111, July 1999. PMID: 10501620.

[301] L F Le Jambre, S Dominik, S J Eady, J M Henshall, and I G Colditz. Adjusting worm egg counts for faecal moisture in sheep. *Veterinary parasitology*, 145(1-2):108–115, April 2007. PMID: 17188813.

[302] D M Leathwick, W E Pomroy, and A C Heath. Anthelmintic resistance in new zealand. *New Zealand Veterinary Journal*, 49(6):227–235, December 2001. PMID: 16032197.

[303] D M Leathwick, T S Waghorn, C M Miller, D S Atkinson, N A Haack, and A-M Oliver. Selective and on-demand drenching of lambs: impact on parasite populations and performance of lambs. *New Zealand veterinary journal*, 54(6):305–312, December 2006. PMID: 17151729.

[304] C. Y. Lee, K. A. Munyard, K. Gregg, J. D. Wetherall, M. J. Stear, and D. M. Groth. The influence of MHC and immunoglobulins a and e on host resistance to gastrointestinal nematodes in sheep. *Journal of Parasitology Research*, 2011:1–11, 2011.

[305] Sang Hong Lee and Julius H J van der Werf. The efficiency of designs for fine-mapping of quantitative trait loci using combined linkage disequilibrium and linkage. *Genetics, selection, evolution: GSE*, 36(2):145–161, April 2004. PMID: 15040896.

[306] Pierre-Charles Lefèvre, Jean Blancou, and René Chermette. *Principales maladies infectieuses et parasitaires du bétail : Europe et régions chaudes*. Éditions Tec & Doc : Éditions médicales internationales, Paris, 2003.

[307] A. Legarra and R. L. Fernando. Linear models for joint association and linkage QTL mapping. *Genet Sel Evol*, 41:43, 2009.

[308] L F LeJambre, R G Windon, and W D Smith. Vaccination against haemonchus contortus: performance of native parasite gut membrane glycoproteins in merino lambs grazing contaminated pasture. *Veterinary parasitology*, 153(3-4):302–312, May 2008. PMID: 18337013.

[309] Gary Leveque, Vincenzo Forgetta, Shaun Morroll, Adrian L Smith, Nat Bumstead, Paul Barrow, J C Loredo-Osti, Kenneth Morgan, and Danielle Malo. Allelic variation in TLR4 is linked to susceptibility to salmonella enterica serovar typhimurium infection in chickens. *Infection and immunity*, 71(3):1116–1124, March 2003. PMID: 12595422.

[310] R C Lewontin. The interaction of selection and linkage. i. general considerations; heterotic models. *Genetics*, 49(1):49–67, January 1964. PMID: 17248194.

[311] R. C. Lewontin. On measures of gametic disequilibrium. *Genetics*, 120(3):849–52, 1988.

[312] R C Lewontin and J Krakauer. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, 74(1):175–195, 1973. PMID: 4711903.

[313] Robert W Li, Sitao Wu, Weizhong Li, Ying Huang, and Louis C Gasbarre. Metagenome plasticity of the bovine abomasal microbiota in immune animals in response to ostertagia ostertagi infection. *PloS one*, 6(9):e24417, 2011. PMID: 21931709.

[314] Robert W Li, Sitao Wu, Weizhong Li, Karl Navarro, Robin D Couch, Dolores Hill, and Jr Urban, Joseph F. Alterations in the porcine colon microbiota induced by the gastrointestinal nematode trichuris suis. *Infection and immunity*, 80(6):2150–2157, June 2012. PMID: 22493085.

[315] Y Li, J E Miller, and D E Franke. Epidemiological observations and heterosis analysis of gastrointestinal nematode parasitism in suffolk, gulf coast native, and crossbred lambs. *Veterinary parasitology*, 98(4):273–283, July 2001. PMID: 11423185.

[316] E. Liénard, G Foucras, F Prévot, C. Grisez, J. P. Bergeaud, D. François, F. Bouvier, and P. Jacquiet. Comparison of gene expression profiles between resistant and susceptible ovine breeds to haemonchus contortus. Buenos Aires, Argentina, 2011.

[317] Peter R. Little, Andrew Hodge, Steven J. Maeder, Nicole C. Wirtherle, David R. Nicholas, George G. Cox, and George A. Conder. Efficacy of a combined oral formulation of derquantel–abamectin against the adult and larval stages of nematodes in sheep, including anthelmintic-resistant strains. *Veterinary Parasitology*, 181(2-4):180–193, September 2011.

[318] S.M. Liu, T.L. Smith, J. Briegel, A. Murray, D.G. Masters, L.J.E. Karlsson, D.G. Palmer, J.C. Greeff, R.B. Besier, and S.B. Gao. Comparing productive performance of nematode resistant merino sheep with non-selected control. *Livestock Production Science*, 97(2-3):117–129, November 2005.

[319] K J Livak and T D Schmittgen. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-delta delta C(T)) method. *Methods (San Diego, Calif.)*, 25(4):402–408, December 2001. PMID: 11846609.

[320] Curtis M. Lively, Clark Craddock, and Robert C. Vrijenhoek. Red queen hypothesis supported by parasitism in sexual and clonal fish. *Nature*, 344(6269):864–866, April 1990.

[321] N Long, D Gianola, G J M Rosa, K A Weigel, and S Avendaño. Machine learning classification procedure for selecting SNPs in genomic selection: application to early mortality in broilers. *Journal of animal breeding and genetics = Zeitschrift für Tierzüchtung und Züchtungsbiologie*, 124(6):377–389, December 2007. PMID: 18076475.

[322] Isis Ludwig-Portugall and Laura E Layland. TLRs, treg, and b cells, an interplay of regulation during helminth infection. *Frontiers in immunology*, 3:8, 2012. PMID: 22566894.

[323] G. Luffau, J.V.T. Khang, J. Bouix, T.C. Nguyen, P. Cullen, and G. Ricordeau. Resistance to experimental infections with haemonchus contortus in romanov sheep. *Genet Sel Evol*, 22:205–229, 1990.

[324] J. P. Lumaret and F. Errouissi. Use of anthelmintics in herbivores and evaluation of risks for the non target fauna of pastures. *Vet Res*, 33(5):547–62, 2002.

[325] M Lynch and B Walsh. *Genetics and analysis of quantitative traits*. Sinauer Assoc., Sunderland, Mass., 1998.

[326] S Maciel, A M Giménez, C Gaona, P J Waller, and J W Hansen. The prevalence of anthelmintic resistance in nematode parasites of sheep in southern latin america: Paraguay. *Veterinary Parasitology*, 62(3-4):207–212, April 1996. PMID: 8686166.

[327] K.M. MacKinnon, J.L. Burton, A.M. Zajac, and D.R. Notter. Microarray analysis reveals difference in gene expression profiles of hair and wool sheep infected with haemonchus contortus. *Veterinary Immunology and Immunopathology*, 130(3-4):210–220, August 2009.

[328] I M MacLeod, B J Hayes, K W Savin, A J Chamberlain, H C McPartlan, and M E Goddard. Power of a genome scan to detect and locate quantitative trait loci in cattle using dense single nucleotide polymorphisms. *Journal of animal breeding and genetics = Zeitschrift für Tierzüchtung und Züchtungsbiologie*, 127(2):133–142, 2010. PMID: 20433522.

[329] Maurice Mahieu, Rémy Arquet, Tony Kandassamy, Nathalie Mandonnet, and Hervé Hoste. Evaluation of targeted drenching using famacha© method in creole goat: Reduction of anthelmintic use, and effects on kid production and pasture contamination. *Veterinary Parasitology*, 146(1-2):135–147, May 2007.

[330] N. Maingi. Resistance to thiabendazole, fenbendazole and levamisole in haemonchus and trichostrongylus species in sheep on a kenyan farm. *Veterinary Parasitology*, 39(3-4):285–291, August 1991.

[331] R. M. Maizels, A. Balic, N. Gomez-Escobar, M. Nair, M. D. Taylor, and J. E. Allen. Helminth parasites–masters of regulation. *Immunol Rev*, 201:89–116, 2004.

[332] R. M. Maizels, E. J. Pearce, D. Artis, M. Yazdanbakhsh, and T. A. Wynn. Regulation of pathogenesis and immunity in helminth infections. *Journal of Experimental Medicine*, 206(10):2059–2066, September 2009.

[333] F S Malan and J A Van Wyk. The packed cell volume and colour of the conjunctivae as aids for monitoring haemonchus contortus infestations in sheep. In *Proc. S. Afr. Vet. Assoc.*, page 139, Grahamstown, 1992.

[334] John H Malone and Brian Oliver. Microarrays, deep sequencing and the true measure of the transcriptome. *BMC biology*, 9:34, 2011. PMID: 21627854.

[335] B. Mangin, B. Goffinet, and A. Rebai. Constructing confidence intervals for QTL location. *Genetics*, 138(4):1301–8, 1994.

[336] K Marshall, J F Maddox, S H Lee, Y Zhang, L Kahn, H-U Graser, C Gondro, S W Walkden-Brown, and J H J van der Werf. Genetic mapping of quantitative trait loci for resistance to haemonchus contortus in sheep. *Animal Genetics*, 40(3):262–272, June 2009. PMID: 19291139.

[337] P J Martin, N Anderson, T Lwin, G Nelson, and T E Morgan. The association between frequency of thiabendazole treatment and the development of resistance in field isolates of ostertagia spp. of sheep. *International Journal for Parasitology*, 14(2):177–181, April 1984. PMID: 6735581.

[338] P J Martin, L F Le Jambre, and J H Claxton. The impact of refugia on the development of thiabendazole resistance in haemonchus contortus. *International Journal for Parasitology*, 11(1):35–41, February 1981. PMID: 7228476.

[339] O. Martinez and R.N. Curnow. Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. *Theoretical and Applied Genetics*, 85(4), December 1992.

[340] M Martínez-Valladares, M P Vara-Del Río, M A Cruz-Rojo, and F A Rojo-Vázquez. Genetic resistance to teladorsagia circumcincta: IgA and parameters at slaughter in churra sheep. *Parasite immunology*, 27(6):213–218, June 2005. PMID: 16048640.

[341] O. Matika, S. Nyoni, J.B. van Wyk, G.J. Erasmus, and R.L. Baker. Resistance of sabi and dorper ewes to gastro-intestinal nematode infections in an african semi-arid environment. *Small Ruminant Research*, 47(2):95–102, February 2003.

[342] O. Matika, R. Pong-Wong, J.A. Wooliams, and S. C. Bishop. Confirmation of two quantitative trait loci regions for nematode resistance in commercial british terminal sire breeds. *Animal*, 5(8):1149–1156, 2011.

[343] Polly Matzinger. The danger model: a renewed sense of self. *Science (New York, N.Y.)*, 296(5566):301–305, April 2002. PMID: 11951032.

[344] A.H. Mbah, A.F. Ogunrinade, and O.A. Dina. Benzimidazole resistance in strongyles of sheep in ibadan, nigeria. *Afr. Livestock Res.*, 1:16–17, 1992.

[345] Mark I McCarthy, Gonçalo R Abecasis, Lon R Cardon, David B Goldstein, Julian Little, John P A Ioannidis, and Joel N Hirschhorn. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature reviews. Genetics*, 9(5):356–369, May 2008. PMID: 18398418.

[346] J. R. McDermott. Mast cells disrupt epithelial barrier function during enteric nematode infection. *Proceedings of the National Academy of Sciences*, 100(13):7761–7766, June 2003.

[347] J H McDonald and M Kreitman. Adaptive protein evolution at the adh locus in drosophila. *Nature*, 351(6328):652–654, June 1991. PMID: 1904993.

[348] L.J. McGaw and J.N. Eloff. Ethnoveterinary use of southern african plants and scientific evaluation of their medicinal properties. *Journal of Ethnopharmacology*, 119(3):559–574, October 2008.

[349] Q A McKellar. Ecotoxicology and residues of anthelmintic compounds. *Veterinary parasitology*, 72(3-4):413–426; discussion 426–435, November 1997. PMID: 9460209.

[350] Q. A. McKellar and F. Jackson. Veterinary anthelmintics: old and new. *Trends in Parasitology*, 20(10):456–461, October 2004.

[351] P B McKenna. The diagnosis value and interpretation of faecal egg counts in sheep. *New Zealand veterinary journal*, 29(8):129–132, August 1981. PMID: 6945520.

[352] R S McLeod. Costs of major parasites to the australian livestock industries. *International Journal for Parasitology*, 25(11):1363–1367, November 1995. PMID: 8635886.

[353] A. F. McRae, J. C. McEwan, K. G. Dodds, T. Wilson, A. M. Crawford, and J. Slate. Linkage disequilibrium in domestic sheep. *Genetics*, 160(3):1113–22, 2002.

[354] J. R. Meadows, E. K. Chan, and J. W. Kijas. Linkage disequilibrium compared between five populations of domestic sheep. *BMC Genet*, 9:61, 2008.

[355] E. N. Meeusen and A. Balic. Do eosinophils have a role in the killing of helminth parasites? *Parasitol Today*, 16(3):95–101, 2000.

[356] Els N T Meeusen, Adam Balic, and Vernon Bowles. Cells, cytokines and other molecules associated with rejection of gastrointestinal nematode parasites. *Veterinary Immunology and Immunopathology*, 108(1-2):121–125, October 2005. PMID: 16099054.

[357] Els N T Meeusen and David Piedrafita. Exploiting natural immunity to helminth parasites for the development of veterinary vaccines. *International journal for parasitology*, 33(11):1285–1290, September 2003. PMID: 13678643.

[358] David M Menge, Jerzy M Behnke, Anna Lowe, John P Gibson, Fuad A Iraqi, R Leyden Baker, and Derek Wakelin. Mapping of chromosomal regions influencing immunological responses to gastrointestinal nematode infections in mice. *Parasite immunology*, 25(6):341–349, June 2003. PMID: 14507332.

[359] M Menzies, A Reverter, N Andronicos, P Hunt, R Windon, and A Ingham. Nematode challenge induces differential expression of oxidant, antioxidant and mucous genes down the longitudinal axis of the sheep gut. *Parasite immunology*, 32(1):36–46, January 2010. PMID: 20042006.

[360] Ted H M Mes, Maarten Eysker, and Harm W Ploeger. A simple, robust and semi-automated parasite egg isolation protocol. *Nature protocols*, 2(3):486–489, 2007. PMID: 17406611.

[361] S.A. Meszaros, J.M. Henshall, S.K. Burgess, G.D. Gray, and B. Tier. Detection of a quantitative trait locus associated with a reduction of faecal egg count in merino sheep. In *Proc. Assoc. Adv. Anim. Breed. Genet.*, volume 13, pages 211–218, 1999.

[362] T H Meuwissen, B J Hayes, and M E Goddard. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157(4):1819–1829, April 2001. PMID: 11290733.

[363] The Meuwissen and Me Goddard. The use of marker haplotypes in animal breeding schemes. *Genetics Selection Evolution*, 28(2):161, 1996.

[364] Theo. Meuwissen. Genomic selection : marker assisted selection on a genome wide scale. *Journal of Animal Breeding and Genetics*, 124(6):321–322, December 2007.

[365] Theo H E Meuwissen and Mike E Goddard. Multipoint identity-by-descent prediction using dense markers to map quantitative trait loci and estimate effective population size. *Genetics*, 176(4):2551–2560, August 2007. PMID: 17565953.

[366] Theo H E Meuwissen, Astrid Karlsen, Sigbjørn Lien, Ingrid Olsaker, and Mike E Goddard. Fine mapping of a quantitative trait locus for twinning rate using combined linkage and linkage disequilibrium mapping. *Genetics*, 161(1):373–379, May 2002. PMID: 12019251.

[367] J F Michel. The epidemiology and control of some nematode infections of grazing animals. *Advances in Parasitology*, 7:211–282, 1969. PMID: 4935269.

[368] J F Michel. The epidemiology and control of some nematode infections in grazing animals. *Advances in Parasitology*, 14:355–397, 1976. PMID: 130791.

[369] J E Miller, M Bahirathan, S L Lemarie, F G Hembry, M T Kearney, and S R Barras. Epidemiology of gastrointestinal nematode parasitism in suffolk and gulf coast native sheep with special emphasis on relative susceptibility to haemonchus contortus infection. *Veterinary parasitology*, 74(1):55–74, January 1998. PMID: 9493310.

[370] J E Miller, S C Bishop, N E Cockett, and R A McGraw. Segregation of natural and experimental gastrointestinal nematode infection in f2 progeny of susceptible suffolk and resistant gulf coast native sheep and its usefulness in assessment of genetic variation. *Veterinary parasitology*, 140(1-2):83–89, August 2006. PMID: 16621290.

[371] J E Miller and D W Horohov. Immunological aspects of nematode parasite control in sheep. *Journal of animal science*, 84 Suppl:E124–132, April 2006. PMID: 16582083.

[372] J. M. Miller, J. Poissant, J. W. Kijas, and D. W. Coltman. A genome-wide set of SNPs detects population substructure and long range linkage disequilibrium in wild sheep. *Mol Ecol Resour*, 11(2):314–22, 2011.

[373] Mohammad Hossein Moradi, Ardeshir Nejati-Javaremi, Mohammad Moradi-Shahrbabak, Ken G Dodds, and John C McEwan. Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. *BMC Genetics*, 13(1):10, 2012.

[374] C. R. Moreno, L. Gruner, A Scala, L. Mura, L Schibler, Y. Amigues, T. Sechi, P. Jacquiet, D. François, S. Sechi, A. Roig, S. Casu, F. Barillet, J. C. Brunel, J. Bouix, A Carta, and R. Rupp. QTL for resistance to internal parasites in two designs based on natural and experimental conditions of infection. volume 15-05, Belo Horizonte, Brazil, 2006.

[375] C. R. Moreno, B. Servin, T Faraut, G. Tosser-Klopp, R. Rupp, P. Mulsant, C Robert-Granié, F. Barillet, C. Delmas, J. Bouix, D. François, D. Allain, E Manfredi, L. Bodin, J M Elsen, D. Robelin, B. Mangin, M.R. Aurel, F. Bouvier, D. Calavas, M. San Cristobal, P. Jacquiet, G Foucras, A. Boissy, and A. Legarra. SheepSNPQTL: utilisation d'une puce 60,000 SNP pour cartographier finement des QTL affectant des caractères de production, de résistance aux maladies et de comportement chez les ovins. volume 16, page 420, Paris, 2009.

[376] E R Morgan, L Cavill, G E Curry, R M Wood, and E S E Mitchell. Effects of aggregation and sample size on composite faecal egg counts in sheep. *Veterinary parasitology*, 131(1-2):79–87, July 2005. PMID: 15921855.

[377] J M Mugambi, R K Bain, S W Wanyangu, M A Ihiga, J L Duncan, M Murray, and M J Stear. Resistance of four sheep breeds to natural and subsequent artificial haemonchus contortus infection. *Veterinary Parasitology*, 69(3-4):265–273, May 1997. PMID: 9195736.

[378] J M Mugambi, S W Wanyangu, R K Bain, M O Owango, J L Duncan, and M J Stear. Response of dorper and red maasai lambs to trickle haemonchus contortus infections. *Research in Veterinary Science*, 61(3):218–221, November 1996. PMID: 8938850.

[379] W M Muir. Comparison of genomic and traditional BLUP-estimated breeding value accuracy and selection response under alternative trait and genomic parameters. *Journal of animal breeding and genetics = Zeitschrift für Tierzüchtung und Züchtungsbiologie*, 124(6):342–355, 2007. PMID: 18076471.

[380] P J Mylrea and I K Hotson. Serum pepsinogen activity and the diagnosis of bovine ostertagiasis. *The British veterinary journal*, 125(8):379–388, August 1969. PMID: 4896805.

[381] A Nari, J Salles, A Gil, P J Waller, and J W Hansen. The prevalence of anthelmintic resistance in nematode parasites of sheep in southern latin america: Uruguay. *Veterinary Parasitology*, 62(3-4):213–222, April 1996. PMID: 8686167.

[382] K.J.N. Ndamukong and M.M.H. Sewell. Resistance to benzimidazole anthelmintics by trichostrongyles in sheep and goats in north-west cameroon. *Veterinary Parasitology*, 41(3-4):335–339, March 1992.

[383] M. Nei. Modification of linkage intensity by natural selection. *Genetics*, 57(3):625–41, 1967.

[384] M. Nei and W. H. Li. Linkage disequilibrium in subdivided populations. *Genetics*, 75(1):213–9, 1973.

[385] M. Nei and W. H. Li. Non-random association between electromorphs and inversion chromosomes in finite populations. *Genet Res*, 35(1):65–83, 1980.

[386] A. Neimann-Sorensen and A. Robertson. The association between blood groups and several production characteristics in three danish cattle breeds. *Acta Agriculturae Scandinavica*, 11(2):163–196, January 1961.

[387] S. E. Newton and E. N. T. Meeusen. Progress and new technologies for developing vaccines against gastrointestinal nematode parasites of sheep. *Parasite Immunology*, 25(5):283–296, May 2003.

[388] S E Newton and E A Munn. The development of vaccines against gastrointestinal nematode parasites, particularly haemonchus contortus. *Parasitology today (Personal ed.)*, 15(3):116–122, March 1999. PMID: 10322325.

[389] R Nielsen. Statistical tests of selective neutrality in the age of genomics. *Heredity*, 86(Pt 6):641–647, June 2001. PMID: 11595044.

[390] R Nielsen. Molecular signatures of natural selection. *Annual Review of Genetics*, 39(1):197–218, December 2005.

[391] G. J. Nieuwhof and S. C. Bishop. Costs of the major endemic diseases of sheep in great britain and the potential benefits of reduction in disease impact. *Animal Science*, 81(01), March 2007.

[392] C Nimbkar, P. M. Ghalsasi, S W Walkden-Brown, and M.N Khan. Evaluation of growth rates and resistance to nematodes of deccani and bannur lambs and their crosses with garole. *Animal*, 73:503–515, 2003.

[393] A J Nisbet, D P Knox, C M McNair, L I Meikle, S K Smith, L A Wildblood, and J B Matthews. Immune recognition of the surface associated antigen, tc-SAA-1, from infective larvae of teladorsagia circumcincta. *Parasite immunology*, 31(1):32–40, January 2009. PMID: 19121081.

[394] D. J. O'Brien, K.L. Strickland, S.L. Parr, J.S. Gray, P.F. O'Reilly, L. Clements, L.M. Dwyer, L.M. O'Reilly, G.C. Coles, F.H.M. Borgsteede, and S. Geerts. Anthelmintic resistance nematodes in farm animals in ireland. In *Anthelmintic resistance in nematodes of farm animals*, pages 101–106, Brussels, November 1994. G.C. Coles, F.H.M. Borgsteede (Eds).

[395] L. O'Connor, S. Walkden-Brown, and L.P. Kahn. Ecology of the free-living stages of major trichostrongylid parasites of sheep. *Veterinary Parasitology*, 142(1-2):1–15, November 2006. PMID: 17011129.

[396] R A Ogunsusi and M Eysker. Inhibited development of trichostrongylids of sheep in northern nigeria. *Research in Veterinary Science*, 26(1):108–110, January 1979. PMID: 472478.

[397] T. K. Oleksyk, M. W. Smith, and S. J. O'Brien. Genome-wide scans for footprints of natural selection. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1537):185–205, November 2009.

[398] Marta I Oliveira, Carine M Gonçalves, Mafalda Pinto, Stéphanie Fabre, Ana Mafalda Santos, Simon F Lee, Mónica A A Castro, Raquel J Nunes, Rita R Barbosa, Jane R Parnes, Chao Yu, Simon J Davis, Alexandra Moreira, Georges Bismuth, and Alexandre M Carmo. CD6 attenuates early and late signaling events, setting thresholds for t-cell activation. *European journal of immunology*, 42(1):195–205, January 2012. PMID: 21956609.

[399] JohanW. Ooijen. Accuracy of mapping quantitative trait loci in autogamous species. *Theoretical and Applied Genetics*, 84-84(7-8), September 1992.

[400] P.M. Outteridge, R.G. Windon, and J.K. Dineen. An association between a lymphocyte antigen in sheep and the response to vaccination against the parasite trichostrongylus colubriformis. *International Journal for Parasitology*, 15(2):121–127, April 1985.

[401] P.M. Outteridge, R.G. Windon, and J.K. Dineen. An ovine lymphocyte antigen marker for acquired resistance to trichostrongylus colubriformis. *International Journal for Parasitology*, 18(6):853–858, September 1988.

[402] P.M. Outteridge, R.G. Windon, J.K. Dineen, and E.F. Smith. The relationship between ovine lymphocyte antigens and faecal egg count of sheep selected for responsiveness to

vaccination against trichostrongylus colubriformis. *International Journal for Parasitology*, 16(4):369–374, August 1986.

[403] D J Overend, M L Phillips, A L Poulton, and C E Foster. Anthelmintic resistance in australian sheep nematode populations. *Australian Veterinary Journal*, 71(4):117–121, April 1994. PMID: 8048907.

[404] Fatih Ozsolak and Patrice M. Milos. RNA sequencing: advances, challenges and opportunities. *Nature Reviews Genetics*, 12(2):87–98, December 2010.

[405] Maria B Padua, Vincent J Lynch, Natalia V Alvarez, Mark A Garthwaite, Thaddeus G Golos, Fuller W Bazer, Satyan Kalkunte, Surendra Sharma, Gunter P Wagner, and Peter J Hansen. ACP5 (Uteroferrin): phylogeny of an ancient and conserved gene expressed in the endometrium of mammals. *Biology of reproduction*, 86(4):123, April 2012. PMID: 22278982.

[406] E Papadopoulos, C Himonas, and G C Coles. Drought and flock isolation may enhance the development of anthelmintic resistance in nematodes. *Veterinary parasitology*, 97(4):253–259, June 2001. PMID: 11390078.

[407] S Paterson. Evidence for balancing selection at the major histocompatibility complex in a free-living ruminant. *The Journal of heredity*, 89(4):289–294, August 1998. PMID: 9703684.

[408] S Paterson, K Wilson, and J M Pemberton. Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population. *Proceedings of the National Academy of Sciences of the United States of America*, 95(7):3714–3719, March 1998. PMID: 9520432.

[409] Tucker A Patterson, Edward K Lobenhofer, Stephanie B Fulmer-Smentek, Patrick J Collins, Tzu-Ming Chu, Wenjun Bao, Hong Fang, Ernest S Kawasaki, Janet Hager, Irina R Tikhonova, Stephen J Walker, Liang Zhang, Patrick Hurban, Francoise de Longueville, James C Fuscoe, Weida Tong, Leming Shi, and Russell D Wolfinger. Performance comparison of one-color and two-color platforms within the microarray quality control (MAQC) project. *Nature Biotechnology*, 24(9):1140–1150, September 2006.

[410] F. Payne. The effect of artificial selection on bristle number in drosophila ampelophila and its interpretation. *Proc Natl Acad Sci U S A*, 4(3):55–8, 1918.

[411] M T Peña, J E Miller, and D W Horohov. Effect of CD4+ t lymphocyte depletion on resistance of gulf coast native lambs to haemonchus contortus infection. *Veterinary parasitology*, 138(3-4):240–246, June 2006. PMID: 16516389.

[412] Alan D Pemberton, Jeremy K Brown, Nicky M Craig, Judith Pate, Kevin McLean, Neil F Inglis, David Knox, and Pamela A Knight. Changes in protein expression in the sheep abomasum following trickle infection with teladorsagia circumcincta. *Parasitology*, 139(3):375–385, March 2012. PMID: 22075947.

[413] Tiago V Pereira, Nikolaos A Patsopoulos, Georgia Salanti, and John P A Ioannidis. Discovery properties of genome-wide association signals from cumulatively combined data sets. *American journal of epidemiology*, 170(10):1197–1206, November 2009. PMID: 19808636.

[414] Jacqueline G Perrigoue, Fraser A Marshall, and David Artis. On the hunt for helminths: innate immune cells in the recognition and response to helminth parasites. *Cellular microbiology*, 10(9):1757–1764, September 2008. PMID: 18505479.

[415] B. D Perry, Robert R Woolaston, G. D Gray, L. F Le Jambre, M. R Knox, B. T Eaton, Jørgen Hansen, R. L Baker, R. A Sani, Roger Carl Merkel, {Subandriyo}, and {Australian Centre for International Agricultural Research}. *Sustainable control of helminth parasites in the tropics*. Australian Centre for International Agricultural Research, Canberra [A.C.T.], 2004.

[416] David Piedrafita, Herman W Raadsma, Jorge Gonzalez, and Els Meeusen. Increased production through parasite control: can ancient breeds of sheep teach us new lessons? *Trends in Parasitology*, 26(12):568–573, December 2010. PMID: 20833589.

[417] Alexander Platt, Bjarni J Vilhjálmsson, and Magnus Nordborg. Conditions under which genome-wide association studies will be positively misleading. *Genetics*, 186(3):1045–1052, November 2010. PMID: 20813880.

[418] Jesse A Poland, Peter J Balint-Kurti, Randall J Wisser, Richard C Pratt, and Rebecca J Nelson. Shades of gray: the world of quantitative disease resistance. *Trends in plant science*, 14(1):21–29, January 2009. PMID: 19062327.

[419] G E Pollott, L J E Karlsson, S Eady, and J C Greeff. Genetic parameters for indicators of host resistance to parasites from weaning to hogget age in merino sheep. *Journal of Animal Science*, 82(10):2852–2864, October 2004. PMID: 15484935.

[420] Alkes L Price, Nick J Patterson, Robert M Plenge, Michael E Weinblatt, Nancy A Shadick, and David Reich. Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics*, 38(8):904–909, 2006. PMID: 16862161.

[421] R. K. Prichard, C. A. Hall, J. D. Kelly, I. C. A. Martin, and A. D. Donald. The problem of anthelmintic resistance in nematodes. *Australian Veterinary Journal*, 56(5):239–250, May 1980.

[422] J. K. Pritchard and M. Przeworski. Linkage disequilibrium in humans: models and data. *Am J Hum Genet*, 69(1):1–14, 2001.

[423] J. K. Pritchard and N. A. Rosenberg. Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet*, 65(1):220–8, 1999.

[424] J K Pritchard, M Stephens, and P Donnelly. Inference of population structure using multilocus genotype data. *Genetics*, 155(2):945–959, June 2000. PMID: 10835412.

[425] J K Pritchard, M Stephens, N A Rosenberg, and P Donnelly. Association mapping in structured populations. *American journal of human genetics*, 67(1):170–181, July 2000. PMID: 10827107.

[426] Franck Prugnolle, Andrea Manica, Marie Charpentier, Jean François Guégan, Vanina Guernier, and François Balloux. Pathogen-driven selection and worldwide HLA class i diversity. *Current biology: CB*, 15(11):1022–1027, June 2005. PMID: 15936272.

[427] M A Rainbird, D Macmillan, and E N Meeusen. Eosinophil-mediated killing of haemonchus contortus larvae: effect of eosinophil activation and role of antibody, complement and interleukin-5. *Parasite immunology*, 20(2):93–103, February 1998. PMID: 9572052.

[428] J.P. Raynaud. Etude de l'efficacité d'une technique de coproscopie quantitative pour le diagnostic de routine et le contrôle des infestations parasitaires des bovins, ovins, équins et porcins. *Ann. Parasitol.*, 45:pp. 321–342, 1970.

[429] Lars Råberg, Andrea L Graham, and Andrew F Read. Decomposing health: tolerance and resistance to parasites in animals. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364(1513):37–49, January 2009. PMID: 18926971.

[430] A Rebai, B Goffinet, and B Mangin. Comparing power of different methods for QTL detection. *Biometrics*, 51(1):87–99, March 1995. PMID: 7766798.

[431] J.E.O Rege, S Tembely, E Mukasa-Mugerwa, S Sovani, D Anindo, A Lahlou-Kassi, S Nagda, and R.L Baker. Effect of breed and season on production and response to infections with gastro-intestinal nematode parasites in sheep in the highlands of ethiopia. *Livestock Production Science*, 78(2):159–174, December 2002.

[432] D E Reich, M Cargill, S Bolk, J Ireland, P C Sabeti, D J Richter, T Lavery, R Kouyoumjian, S F Farhadian, R Ward, and E S Lander. Linkage disequilibrium in the human genome. *Nature*, 411(6834):199–204, May 2001. PMID: 11346797.

[433] J A Requejo-Fernández, A Martínez, A Meana, F A Rojo-Vázquez, K Osoro, and L M Ortega-Mora. Anthelmintic resistance in nematode parasites from goats in spain. *Veterinary Parasitology*, 73(1-2):83–88, December 1997. PMID: 9477495.

[434] 3rd Rexroad, Caird E and Roger L Vallejo. Estimates of linkage disequilibrium and effective population size in rainbow trout. *BMC Genetics*, 10:83, 2009. PMID: 20003428.

[435] D G Riley and J A Van Wyk. Genetic parameters for FAMACHA score and related traits for host resistance/resilience and production at differing severities of worm challenge in a merino flock in south africa. *Veterinary parasitology*, 164(1):44–52, September 2009. PMID: 19446960.

[436] J Riquet, W Coppieters, N Cambisano, J J Arranz, P Berzi, S K Davis, B Grisart, F Farnir, L Karim, M Mni, P Simon, J F Taylor, P Vanmanshoven, D Wagenaar, J E Womack, and M Georges. Fine-mapping of quantitative trait loci by identity by descent in outbred populations: application to milk production in dairy cattle. *Proceedings of the National Academy of Sciences of the United States of America*, 96(16):9252–9257, August 1999. PMID: 10430929.

[437] Nicholas Robinson, David Piedrafita, Kenneth Snibson, Paul Harrison, and Els N Meeusen. Immune cell kinetics in the ovine abomasal mucosa following hyperimmunization and

challenge with haemonchus contortus. *Veterinary Research*, 41(4):37, August 2010. PMID: 20167198.

[438] Nicholas Robinson, Jill Pleasance, David Piedrafita, and Els N Meeusen. The kinetics of local cytokine and galectin expression after challenge infection with the gastrointestinal nematode, haemonchus contortus. *International Journal for Parasitology*, 41(5):487–493, April 2011. PMID: 21172352.

[439] O. W. Robison, B. T. McDaniel, and E. J. Rincon. Estimation of direct and maternal additive and heterotic effects from crossbreeding experiments in animals. *Journal of Animal Science*, 52:44–50, 1981.

[440] Matthew V Rockman. Reverse engineering the genotype-phenotype map with natural genetic variation. *Nature*, 456(7223):738–744, December 2008. PMID: 19079051.

[441] E Romjali, P Dorny, A Batubara, V S Pandey, and R M Gatenby. Peri-parturient rise in faecal strongyle egg counts of different genotypes of sheep in north sumatra, indonesia. *Veterinary parasitology*, 68(1-2):191–196, January 1997. PMID: 9066065.

[442] M Ron and J I Weller. From QTL to QTN identification in livestock–winning by points rather than knock-out: a review. *Animal Genetics*, 38(5):429–439, October 2007. PMID: 17697134.

[443] Marc E Rothenberg and Simon P Hogan. The eosinophil. *Annual review of immunology*, 24:147–174, 2006. PMID: 16551246.

[444] A Rowe, C Gondro, D Emery, and N Sangster. Sequential microarray to identify timing of molecular responses to haemonchus contortus infection in sheep. *Veterinary parasitology*, 161(1-2):76–87, April 2009. PMID: 19200661.

[445] A Rowe, K McMaster, D Emery, and N Sangster. Haemonchus contortus infection in sheep: parasite fecundity correlates with worm size and host lymphocyte counts. *Veterinary parasitology*, 153(3-4):285–293, 2008. PMID: 18337014.

[446] R. Rupp, D Boichard, A. Barbat, J.M. Astruc, G. Lagriffoul, and F. Barillet. Selection for mastitis resistance in french dairy sheep. Montpellier, 2002.

[447] Su Youn Ryu, Gwan Ui Hong, Dae Yong Kim, and Jai Youl Ro. Enolase 1 and calreticulin regulate the differentiation and function of mouse mast cells. *Cellular signalling*, 24(1):60–70, January 2012. PMID: 21803152.

[448] Pardis C Sabeti, David E Reich, John M Higgins, Haninah Z P Levine, Daniel J Richter, Stephen F Schaffner, Stacey B Gabriel, Jill V Platko, Nick J Patterson, Gavin J McDonald, Hans C Ackerman, Sarah J Campbell, David Altshuler, Richard Cooper, Dominic Kwiatkowski, Ryk Ward, and Eric S Lander. Detecting recent positive selection in the human genome from haplotype structure. *Nature*, 419(6909):832–837, October 2002. PMID: 12397357.

[449] Pardis C Sabeti, Patrick Varilly, Ben Fry, Jason Lohmueller, Elizabeth Hostetter, Chris Cotsapas, Xiaohui Xie, Elizabeth H Byrne, Steven A McCarroll, Rachelle Gaudet, Stephen F Schaffner, Eric S Lander, Kelly A Frazer, Dennis G Ballinger, David R Cox, David A Hinds, Laura L Stuve, Richard A Gibbs, John W Belmont, Andrew Boudreau, Paul Hardenbol, Suzanne M Leal, Shiran Pasternak, David A Wheeler, Thomas D Willis, Fuli Yu, Huanming Yang, Changqing Zeng, Yang Gao, Haoran Hu, Weitao Hu, Chaohua Li, Wei Lin, Siqi Liu, Hao Pan, Xiaoli Tang, Jian Wang, Wei Wang, Jun Yu, Bo Zhang, Qingrun Zhang, Hongbin Zhao, Hui Zhao, Jun Zhou, Stacey B Gabriel, Rachel Barry, Brendan Blumenstiel, Amy Camargo, Matthew Defelice, Maura Faggart, Mary Goyette, Supriya Gupta, Jamie Moore, Huy Nguyen, Robert C Onofrio, Melissa Parkin, Jessica Roy, Erich Stahl, Ellen Winchester, Liuda Ziaugra, David Altshuler, Yan Shen, Zhijian Yao, Wei Huang, Xun Chu, Yungang He, Li Jin, Yangfan Liu, Yayun Shen, Weiwei Sun, Haifeng Wang, Yi Wang, Ying Wang, Xiaoyan Xiong, Liang Xu, Mary M Y Waye, Stephen K W Tsui, Hong Xue, J Tze-Fei Wong, Luana M Galver, Jian-Bing Fan, Kevin Gunderson, Sarah S Murray, Arnold R Oliphant, Mark S Chee, Alexandre Montpetit, Fanny Chagnon, Vincent Ferretti, Martin Leboeuf, Jean-François Olivier, Michael S Phillips, Stéphanie Roumy, Clémentine Sallée, Andrei Verner, Thomas J Hudson, Pui-Yan Kwok, Dongmei Cai, Daniel C Koboldt, Raymond D Miller, Ludmila Pawlikowska, Patricia Taillon-Miller, Ming Xiao, Lap-Chee Tsui, William Mak, You Qiang Song, Paul K H Tam, Yusuke Nakamura, Takahisa Kawaguchi, Takuya Kitamoto, Takashi Morizono, Atsushi Nagashima, Yozo Ohnishi, Akihiro Sekine, Toshihiro Tanaka, Tatsuhiko Tsunoda, Panos Deloukas, Christine P Bird, Marcos Delgado, Emmanouil T Dermitzakis, Rhian Gwilliam, Sarah

Hunt, Jonathan Morrison, Don Powell, Barbara E Stranger, Pamela Whittaker, David R Bentley, Mark J Daly, Paul I W de Bakker, Jeff Barrett, Yves R Chretien, Julian Maller, Steve McCarroll, Nick Patterson, Itsik Pe'er, Alkes Price, Shaun Purcell, Daniel J Richter, Pardis Sabeti, Richa Saxena, Stephen F Schaffner, Pak C Sham, Patrick Varilly, David Altshuler, Lincoln D Stein, Lalitha Krishnan, Albert Vernon Smith, Marcela K Tello-Ruiz, Gudmundur A Thorisson, Aravinda Chakravarti, Peter E Chen, David J Cutler, Carl S Kashuk, Shin Lin, Gonçalo R Abecasis, Weihua Guan, Yun Li, Heather M Munro, Zhaohui Steve Qin, Daryl J Thomas, Gilean McVean, Adam Auton, Leonardo Bottolo, Niall Cardin, Susana Eyheramendy, Colin Freeman, Jonathan Marchini, Simon Myers, Chris Spencer, Matthew Stephens, Peter Donnelly, Lon R Cardon, Geraldine Clarke, David M Evans, Andrew P Morris, Bruce S Weir, Tatsuhiko Tsunoda, Todd A Johnson, James C Mullikin, Stephen T Sherry, Michael Feolo, Andrew Skol, Houcan Zhang, Changqing Zeng, Hui Zhao, Ichiro Matsuda, Yoshimitsu Fukushima, Darryl R Macer, Eiko Suda, Charles N Rotimi, Clement A Adebamowo, Ike Ajayi, Toyin Aniagwu, Patricia A Marshall, Chibuzor Nkwodimmah, Charmaine D M Royal, Mark F Leppert, Missy Dixon, Andy Peiffer, Renzong Qiu, Alastair Kent, Kazuto Kato, Norio Niikawa, Isaac F Adewole, Bartha M Knoppers, Morris W Foster, Ellen Wright Clayton, Jessica Watkin, Richard A Gibbs, John W Belmont, Donna Muzny, Lynne Nazareth, Erica Sodergren, George M Weinstock, David A Wheeler, Imtaz Yakub, Stacey B Gabriel, Robert C Onofrio, Daniel J Richter, Liuda Ziaugra, Bruce W Birren, Mark J Daly, David Altshuler, Richard K Wilson, Lucinda L Fulton, Jane Rogers, John Burton, Nigel P Carter, Christopher M Clee, Mark Griffiths, Matthew C Jones, Kirsten McLay, Robert W Plumb, Mark T Ross, Sarah K Sims, David L Willey, Zhu Chen, Hua Han, Le Kang, Martin Godbout, John C Wallenburg, Paul L'Archevêque, Guy Bellemare, Koji Saeki, Hongguang Wang, Daochang An, Hongbo Fu, Qing Li, Zhen Wang, Renwu Wang, Arthur L Holden, Lisa D Brooks, Jean E McEwen, Mark S Guyer, Vivian Ota Wang, Jane L Peterson, Michael Shi, Jack Spiegel, Lawrence M Sung, Lynn F Zacharia, Francis S Collins, Karen Kennedy, Ruth Jamieson, and John Stewart. Genome-wide detection and characterization of positive selection in human populations. *Nature*, 449(7164):913–918, October 2007. PMID: 17943131.

[450] D. Sackett, P. Holmes, K. Abbott, S.i Jephcott, and M. Barber. Assessing the economic cost of endemic disease on the profitability of australian beef cattle and sheep producers. Technical report, Meat and Livestock Australia, Australia, 2006.

[451] H A Saddiqi, Z Iqbal, M N Khan, M Sarwar, G Muhammad, M Yaseen, and A Jabbar. Evaluation of three pakistani sheep breeds for their natural resistance to artificial infection of haemonchus contortus. *Veterinary Parasitology*, 168(1-2):141–145, February 2010. PMID: 19939567.

[452] E. Safari, N.M. Fogarty, and A.R. Gilmour. A review of genetic parameter estimates for wool, growth, meat and reproduction traits in sheep. *Livestock Production Science*, 92(3):271–289, March 2005.

[453] G. Sallé, F Guillaume, A. Legarra, and C. R. Moreno. Proposition of procedures to handle SNP data: example of an ovine population genotyped with the OvineSNP50 beadchip. Paris, 2010.

[454] G Sallé, P Jacquiet, L Gruner, J Cortet, C Sauvé, F Prévot, C Grisez, J P Bergeaud, L Schibler, A Tircazes, D François, C Pery, F Bouvier, J C Thouly, J C Brunel, A Legarra, J M Elsen, J Bouix, R Rupp, and C R Moreno. A genome scan for QTL affecting resistance to haemonchus contortus in sheep. *Journal of animal science*, July 2012. PMID: 22767094.

[455] N C Sangster and J Gill. Pharmacology of anthelmintic resistance. *Parasitology Today (Personal Ed.)*, 15(4):141–146, April 1999. PMID: 10322335.

[456] N. D. Sargison. Pharmaceutical control of endoparasitic helminth infections in sheep. *Veterinary Clinics of North America: Food Animal Practice*, 27(1):139–156, March 2011.

[457] N. D. Sargison, F. Jackson, D. J. Bartley, D. J. Wilson, L. J. Stenhouse, and C. D. Penny. Observations on the emergence of multiple anthelmintic resistance in sheep flocks in the south-east of scotland. *Vet Parasitol*, 145(1-2):65–76, 2007.

[458] G Sayers, B Good, J P Hanrahan, M Ryan, J M Angles, and T Sweeney. Major histocompatibility complex DRB1 gene: its role in nematode resistance in suffolk and texel sheep breeds. *Parasitology*, 131(Pt 3):403–409, September 2005. PMID: 16178362.

[459] G Sayers, B Good, J P Hanrahan, M Ryan, and T Sweeney. Intron 1 of the interferon gamma gene: Its role in nematode resistance in suffolk and texel sheep breeds. *Research in veterinary science*, 79(3):191–196, December 2005. PMID: 16054889.

[460] B L Sayre and G C Harris. Systems genetics approach reveals candidate genes for parasite resistance from quantitative trait loci studies in agricultural species. *Animal Genetics*, 43(2):190–198, April 2012. PMID: 22404355.

[461] H D Schallig, M A van Leeuwen, W E Bernadina, and W M Hendrikx. Serum antibody responses of texel sheep experimentally infected with haemonchus contortus. *Research in veterinary science*, 57(1):63–68, July 1994. PMID: 7973095.

[462] H D Schallig, M A van Leeuwen, and A W Cornelissen. Protective immunity induced by vaccination with two haemonchus contortus excretory secretory proteins in sheep. *Parasite immunology*, 19(10):447–453, October 1997. PMID: 9372512.

[463] Jan H Schefe, Kerstin E Lehmann, Ivo R Buschmann, Thomas Unger, and Heiko Funke-Kaiser. Quantitative real-time RT-PCR data analysis: current concepts and the novel "gene expression's CT difference" formula. *Journal of molecular medicine (Berlin, Germany)*, 84(11):901–910, November 2006. PMID: 16972087.

[464] M Schena, D Shalon, R W Davis, and P O Brown. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science (New York, N.Y.)*, 270(5235):467–470, October 1995. PMID: 7569999.

[465] R. D. Schulte, C. Makus, B. Hasert, N. K. Michiels, and H. Schulenburg. Multiple reciprocal adaptations and rapid genetic change upon experimental coevolution of an animal host and its microbial parasite. *Proceedings of the National Academy of Sciences*, 107(16):7359–7364, April 2010.

[466] Rebecca D Schulte, Carsten Makus, Barbara Hasert, Nico K Michiels, and Hinrich Schulenburg. Multiple reciprocal adaptations and rapid genetic change upon experimental coevolution of an animal host and its microbial parasite. *Proceedings of the National Academy of Sciences of the United States of America*, 107(16):7359–7364, April 2010. PMID: 20368449.

[467] F.-W. Schwaiger, D. Gostomski, M.J. Stear, J.L. Duncan, Q.A. McKellar, J.T. Epplen, and J. Buitkamp. An ovine major histocompatibility complex DRB1 allele is associated with low faecal egg counts following natural, predominantly ostertagia circumcincta infection. *International Journal for Parasitology*, 25(7):815–822, July 1995.

[468] R.J. Shaw, C.A. Morris, M. Wheeler, M. Tate, and I.A. Sutherland. Salivary IgA: a suitable measure of immunity to gastrointestinal nematodes in sheep. *Veterinary Parasitology*, 186(1-2):109–117, May 2012.

[469] M V B Silva, T S Sonstegard, O Hanotte, J M Mugambi, J F Garcia, S Nagda, J P Gibson, F A Iraqi, A E McClintock, S J Kemp, P J Boettcher, M Malek, C P Van Tassell, and R L Baker. Identification of quantitative trait loci affecting resistance to gastrointestinal parasites in a double backcross population of red maasai and dorper sheep. *Animal Genetics*, 43(1):63–71, February 2012. PMID: 22221026.

[470] Anne Silvestre, Vincent Leignel, Boumadiane Berrag, Nadine Gasnier, Jean-François Humbert, Christophe Chartiere, and Jacques Cabaret. Sheep and goat nematode resistance to anthelmintics: pro and cons among breeding management factors. *Veterinary Research*, 33(5):465–480, October 2002. PMID: 12387484.

[471] H. V. Simpson. Pathophysiology of abomasal parasitism: is the host or parasite responsible? *Vet J*, 160(3):177–91, 2000.

[472] H. V. Simpson, S. M. Przemeck, I. Scott, D. G. Thomas, R. S. Green, and G. W. Reynolds. Pathophysiology in teladorsagia (Ostertagia) circumcincta-infected sheep selected for high fleece weight. *Vet Parasitol*, 163(1-2):73–80, 2009.

[473] M. M. Sissay, A Asefa, A Uggla, and P. J. Waller. Anthelmintic resistance of nematode parasites of small ruminants in eastern ethiopia: Exploitation of refugia to restore anthelmintic efficacy. *Veterinary Parasitology*, 135(3-4):337–346, February 2006.

[474] M. Slatkin. Linkage disequilibrium in growing and stable populations. *Genetics*, 137(1):331–6, 1994.

[475] M. Slatkin. Linkage disequilibrium–understanding the evolutionary past and mapping the medical future. *Nat Rev Genet*, 9(6):477–85, 2008.

[476] Charles Smith. Improvement of metric traits through specific genetic loci. *Animal Production*, 9(03):349–358, 1967.

[477] W D Smith and D S Zarlenga. Developments and hurdles in generating vaccines for controlling helminth parasites of grazing ruminants. *Veterinary parasitology*, 139(4):347–359, July 2006. PMID: 16750599.

[478] F. Soli, T.H. Terrill, S.A. Shaik, W.R. Getz, J.E. Miller, M. Vanguru, and J.M. Burke. Efficacy of copper oxide wire particles against gastrointestinal nematodes in sheep and goats. *Veterinary Parasitology*, 168(1-2):93–96, February 2010.

[479] R J Spelman and H Bovenhuis. Moving from QTL experimental results to the utilization of QTL in breeding programmes. *Animal genetics*, 29(2):77–84, April 1998. PMID: 9699266.

[480] Dina A St Clair. Quantitative disease resistance and quantitative resistance loci in breeding. *Annual review of phytopathology*, 48:247–268, 2010. PMID: 19400646.

[481] K A Stafford, E R Morgan, and G C Coles. Weight-based targeted selective treatment of gastrointestinal nematodes in a commercial sheep flock. *Veterinary Parasitology*, 164(1):59–65, September 2009. PMID: 19428187.

[482] M J Stear, K Bairden, J L Duncan, P H Holmes, Q A McKellar, M Park, S Strain, M Murray, S C Bishop, and G Gettinby. How hosts control worms. *Nature*, 389(6646):27, September 1997. PMID: 9288962.

[483] M. J. Stear and S. C. Bishop. The curvilinear relationship between worm length and fecundity of teladorsagia circumcincta. *Int J Parasitol*, 29(5):777–80, 1999.

[484] M J Stear, S C Bishop, M Doligalska, J L Duncan, P H Holmes, J Irvine, L McCririe, Q A McKellar, E Sinski, and M Murray. Regulation of egg production, worm burden, worm length and worm fecundity by host responses in sheep infected with ostertagia circumcincta. *Parasite immunology*, 17(12):643–652, December 1995. PMID: 8834764.

[485] M J Stear, S C Bishop, B A Mallard, and H Raadsma. The sustainability, feasibility and desirability of breeding livestock for disease resistance. *Research in veterinary science*, 71(1):1–7, August 2001. PMID: 11666141.

[486] M J Stear, B Boag, I Cattadori, and L Murphy. Genetic variation in resistance to mixed, predominantly teladorsagia circumcincta nematode infections of sheep: from heritabilities to gene identification. *Parasite immunology*, 31(5):274–282, May 2009. PMID: 19388948.

[487] M. J. Stear, M. Doligalska, and K. Donskow-Schmelter. Alternatives to anthelmintics for the control of nematodes in livestock. *Parasitology*, 134(02):139, November 2006.

[488] M J Stear, G T Innocent, and J Buitkamp. The evolution and maintenance of polymorphism in the major histocompatibility complex. *Veterinary immunology and immunopathology*, 108(1-2):53–57, October 2005. PMID: 16099055.

[489] M. J. Stear, M. Park, and S. C. Bishop. The key components of resistance to ostertagia circumcincta in lambs. *Parasitol Today*, 12(11):438–41, 1996.

[490] Michael J Stear, Lesley Fitton, Giles T Innocent, Lisa Murphy, Kerry Rennie, and Louise Matthews. The dynamic influence of genetic variation on the susceptibility of sheep to gastrointestinal nematode infection. *Journal of the Royal Society, Interface / the Royal Society*, 4(16):767–776, October 2007. PMID: 17626002.

[491] J. D. Storey and R. Tibshirani. Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A*, 100(16):9440–5, 2003.

[492] Subandryio. Indonesian thin tail sheep performances on production, reproduction, their resistance to internal parasites, and the efforts to improve their productivity. page 24, Bogor, Indonesia, September 2002.

[493] Robert W Summers, David E Elliott, Khurram Qadir, Jr Urban, Joseph F, Robin Thompson, and Joel V Weinstock. Trichuris suis seems to be safe and possibly effective in the treatment of inflammatory bowel disease. *The American journal of gastroenterology*, 98(9):2034–2041, September 2003. PMID: 14499784.

[494] I. A. Sutherland, D. M. Leathwick, A. E. Brown, and C. M. Miller. Prophylactic efficacy of persistent anthelmintics against challenge with drug-resistant and susceptible ostertagia circumcincta. *Veterinary Record*, 141(5):120–123, August 1997.

[495] Ian A. Sutherland and Dave M. Leathwick. Anthelmintic resistance in nematode parasites of cattle: a global issue? *Trends in Parasitology*, 27(4):176–181, April 2011.

[496] N B Sutter, M A Eberle, Hi G Parker, B J Pullar, E F Kirkness, L Kruglyak, and E A Ostrander. Extensive and breed-specific linkage disequilibrium in canis familiaris. *Genome Research*, 14(12):2388–2396, December 2004. PMID: 15545498.

[497] Joshua P A Sweeny, Ian D Robertson, Una M Ryan, Caroline Jacobson, and Rob G Woodgate. Comparison of molecular and McMaster microscopy techniques to confirm the

presence of naturally acquired strongylid nematode infections in sheep. *Molecular and biochemical parasitology*, 180(1):62–67, November 2011. PMID: 21827796.

[498] J. Tabel. *Alternatives au traitement chimiothérapeutique des strongyloses gastro-intestinales des ovins: bilan et perspectives.* Thèse d'exercice, Ecole Nationale Vétérinaire de Toulouse, ENVT, 2011.

[499] F Tajima. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123(3):585–595, November 1989. PMID: 2513255.

[500] K.A. Tariq, M.Z. Chishti, F. Ahmad, and A.S. Shawl. Epidemiology of gastrointestinal nematodes of sheep managed under traditional husbandry system in kashmir valley. *Veterinary Parasitology*, 158(1-2):138–143, November 2008.

[501] M.A. Taylor. SCOPS and COWS—'Worming it out of UK farmers'. *Veterinary Parasitology*, November 2011.

[502] S Taylor, M. Wakem, G. Dijkman, M. Alsarraj, and M. Nguyen. A practical approach to RT-qPCR—Publishing data that conform to the MIQE guidelines. *Methods*, 50(4):S1–S5, April 2010.

[503] S Tembely, A Lahlou-Kassi, J E Rege, E Mukasa-Mugerwa, D Anindo, S Sovani, and R L Baker. Breed and season effects on the peri-parturient rise in nematode egg output in indigenous ewes in a cool tropical environment. *Veterinary parasitology*, 77(2-3):123–132, June 1998. PMID: 9746282.

[504] A Tenesa, P Navarro, B J Hayes, D L Duffy, G M Clarke, M E Goddard, and P M Visscher. Recent human effective population size estimated from linkage disequilibrium. *Genome Research*, 17(4):520–526, April 2007. PMID: 17351134.

[505] G. Terefe. *Rôle des éosinophiles dans la régulation des populations d'Haemonchus contortus chez le mouton.* PhD thesis, INP, Toulouse, 2007.

[506] G. Terefe, C. Grisez, F. Prevot, J. P. Bergeaud, P. Dorchies, J. C. Brunel, D. Francois, I. Fourquaux, and P. Jacquiet. In vitro pre-exposure of haemonchus contortus l3 to blood eosinophils reduces their establishment potential in sheep. *Vet Res*, 38(4):647–54, 2007.

[507] G. Terefe, C. Lacroux, O. Andreoletti, C. Grisez, F. Prevot, J. P. Bergeaud, J. Penicaud, V. Rouillon, L. Gruner, J. C. Brunel, D. Francois, J. Bouix, P. Dorchies, and P. Jacquiet.

Immune response to haemonchus contortus infection in susceptible (INRA 401) and resistant (Barbados black belly) breeds of lambs. *Parasite Immunol*, 29(8):415–24, 2007.

[508] G Terefe, C Lacroux, F Prévot, C Grisez, J P Bergeaud, C Bleuart, P Dorchies, G Foucras, and P Jacquiet. Eosinophils in haemonchus contortus-infected resistant and susceptible breeds of sheep: abomasal tissue recruitment and in vitro functional state. *Veterinary Parasitology*, 165(1-2):161–164, October 2009. PMID: 19733438.

[509] S Teyssèdre, M C Dupuis, G Guérin, L Schibler, J M Denoix, J M Elsen, and A Ricard. Genome-wide association studies for osteochondrosis in french trotter horses. *Journal of Animal Science*, 90(1):45–53, January 2012. PMID: 21841084.

[510] G. Thomson. The effect of a selected locus on linked neutral loci. *Genetics*, 85(4):753–88, 1977.

[511] S Thévenon, G K Dayo, S Sylla, I Sidibe, D Berthier, H Legros, D Boichard, A Eggen, and M Gautier. The extent of linkage disequilibrium in a large cattle population of western africa and its consequences for association studies. *Animal Genetics*, 38(3):277–286, June 2007. PMID: 17459015.

[512] R Tinar, C V Akyol, V Y Cirak, B Senlik, and C Bauer. Investigations on the seasonal patterns of strongyle infections in grazing lambs, and the occurrence of anthelmintic resistance on sheep and goat farms in western anatolia, turkey. *Parasitology Research*, 96(1):18–23, April 2005. PMID: 15765247.

[513] Ian R Tizard and Ian Rodney Tizard. *Veterinary immunology : an introduction*. W.B. Saunders, Philadelphia [etc.], 1996.

[514] Paul R Torgerson, Michaela Paul, and Fraser I Lewis. The contribution of simple random sampling to observed variations in faecal egg counts. *Veterinary parasitology*, 188(3-4):397–401, September 2012. PMID: 22521975.

[515] A Torina, V Ferrantelli, O A E Sparagano, S Reale, F Vitale, and S Caracappa. Climatic conditions and gastrointestinal nematode egg production: observations in breeding sheep and goats. *Annals of the New York Academy of Sciences*, 1026:203–209, October 2004. PMID: 15604494.

[516] D G Turner, L A Wildblood, N F Inglis, and D G Jones. Characterization of a galectin-like activity from the parasitic nematode, haemonchus contortus, which modulates ovine eosinophil migration in vitro. *Veterinary immunology and immunopathology*, 122(1-2):138–145, March 2008. PMID: 18187208.

[517] E Uleberg, I S Widerøe, E Grindflek, J Szyda, S Lien, and T H E Meuwissen. Fine mapping of a QTL for intramuscular fat on porcine chromosome 6 using combined linkage and linkage disequilibrium mapping. *Journal of animal breeding and genetics = Zeitschrift für Tierzüchtung und Züchtungsbiologie*, 122(1):1–6, February 2005. PMID: 16130482.

[518] D Vagenas, A Doeschl-Wilson, S C Bishop, and I Kyriazakis. In silico exploration of the effects of host genotype and nutrition on the genetic parameters of lambs challenged with gastrointestinal parasites. *International journal for parasitology*, 37(14):1617–1630, December 2007. PMID: 17658533.

[519] E H van der Waaij, P Bijma, S C Bishop, and J A van Arendonk. Modeling selection for production traits under constant infection pressure. *Journal of animal science*, 78(11):2809–2820, November 2000. PMID: 11063303.

[520] J H J van der Werf. Marker assisted selection in sheep and goat industries. In *Marker-assisted selection : current status and future perspectives in crops, livestock, forestry and fish*. Food and Agriculture Organization of the United Nations, Rome, 2007.

[521] L Van Valen. A new evolutionary law. *Evolutionary Theory*, 1(1):1–30, 1973.

[522] J A van Wyk. Refugia–overlooked as perhaps the most potent factor concerning the development of anthelmintic resistance. *The Onderstepoort Journal of Veterinary Research*, 68(1):55–67, March 2001. PMID: 11403431.

[523] J A van Wyk, H Hoste, R M Kaplan, and R B Besier. Targeted selective treatment for worm management–how do we sell rational programs to farmers? *Veterinary Parasitology*, 139(4):336–346, July 2006. PMID: 16774807.

[524] J A Van Wyk, F S Malan, H M Gerber, and R M Alves. The problem of escalating resistance of haemonchus contortus to the modern anthelmintics in south africa. *The Onderstepoort Journal of Veterinary Research*, 56(1):41–49, March 1989. PMID: 2726192.

[525] J A van Wyk, F S Malan, and J L Randles. How long before resistance makes it impossible to control some field strains of haemonchus contortus in south africa with any of the modern anthelmintics? *Veterinary Parasitology*, 70(1-3):111–122, June 1997. PMID: 9195715.

[526] Jan A. Van Wyk and Gareth F. Bath. The FAMACHA system for managing haemonchosis in sheep and goats by clinically identifying individual animals for treatment. *Veterinary Research*, 33(5):509–529, September 2002.

[527] Jo Vandesompele, Katleen De Preter, Filip Pattyn, Bruce Poppe, Nadine Van Roy, Anne De Paepe, and Frank Speleman. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, 3(7):RESEARCH0034, June 2002. PMID: 12184808.

[528] Gerardo R Vasta. Roles of galectins in infection. *Nature reviews. Microbiology*, 7(6):424–438, June 2009. PMID: 19444247.

[529] F Veglia. The anatomy and life history of the haemonchus contortus (Rud). *Ann Rep Dir Vet S Afr*, 3(4):347–500, 1915.

[530] J Vercruysse, P Dorny, and K Meurrens. Benzimidazole resistance of nematodes in sheep in belgium. *The Veterinary Record*, 125(24):602–603, December 1989. PMID: 2609486.

[531] Jozef Vercruysse, David P Knox, Theo P M Schetters, and Peter Willadsen. Veterinary parasitic vaccines: pitfalls and future directions. *Trends in parasitology*, 20(10):488–492, October 2004. PMID: 15363443.

[532] Cecília José Veríssimo, Simone Cristina Méo Niciura, Ana Lúcia Luz Alberti, Carlos Frederico Carvalho Rodrigues, Cristina Maria Pacheco Barbosa, Daniela Pontes Chiebao, Daniel Cardoso, Giane Serafim da Silva, José Roberto Pereira, Luiz Florêncio Franco Margatho, Ricardo Lopes Dias da Costa, Romeu Fernandes Nardon, Tatiana Evelyn Hayama Ueno, Vera Cláudia Lorenzetti Magalhães Curci, and Marcelo Beltrão Molento. Multidrug and multispecies resistance in sheep flocks from são paulo state, brazil. *Veterinary Parasitology*, January 2012. PMID: 22341829.

[533] L Vervelde, M A W Van Leeuwen, M Kruidenier, F N J Kooyman, J F Huntley, I Van Die, and A W C A Cornelissen. Protection studies with recombinant excretory/secretory pro-

teins of haemonchus contortus. *Parasite immunology*, 24(4):189–201, April 2002. PMID: 12010484.

[534] L Viegi, A Pieroni, P Guarrera, and R Vangelisti. A review of plants used in folk veterinary medicine in italy as basis for a databank. *Journal of Ethnopharmacology*, 89(2-3):221–244, December 2003. PMID: 14611886.

[535] Alain Vignal, Denis Milan, Magali SanCristobal, and André Eggen. A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics, Selection, Evolution: GSE*, 34(3):275–305, June 2002. PMID: 12081799.

[536] Beatriz Villanueva, Ricardo Pong-Wong, and John A Woolliams. Marker assisted selection with optimised contributions of the candidates to selection. *Genetics, selection, evolution: GSE*, 34(6):679–703, December 2002. PMID: 12486398.

[537] P. M. Visscher and M. E. Goddard. Prediction of the confidence interval of quantitative trait loci location. *Behav Genet*, 34(4):477–82, 2004.

[538] P. M. Visscher and C. S. Haley. Detection of putative quantitative trait loci in line crosses under infinitesimal genetic models. *Theoretical and Applied Genetics*, 93-93(5-6):691–702, October 1996.

[539] P. M. Visscher, R. Thompson, and C. S. Haley. Confidence intervals in QTL mapping by bootstrapping. *Genetics*, 143(2):1013–20, 1996.

[540] David Voehringer. The role of basophils in helminth infection. *Trends in parasitology*, 25(12):551–556, December 2009. PMID: 19782643.

[541] David Voehringer, N van Rooijen, and Richard M Locksley. Eosinophils develop in distinct stages and are recruited to peripheral sites by alternatively activated macrophages. *Journal of leukocyte biology*, 81(6):1434–1444, June 2007. PMID: 17339609.

[542] P Vos, R Hogers, M Bleeker, M Reijans, T van de Lee, M Hornes, A Frijters, J Pot, J Peleman, and M Kuiper. AFLP: a new technique for DNA fingerprinting. *Nucleic acids research*, 23(21):4407–4414, November 1995. PMID: 7501463.

[543] Seth T Walk, Arthur M Blum, Sarah Ang-Sheng Ewing, Joel V Weinstock, and Vincent B Young. Alteration of the murine gut microbiota during infection with the para-

sitic helminth heligmosomoides polygyrus. *Inflammatory bowel diseases*, 16(11):1841–1849, November 2010. PMID: 20848461.

[544] J. D. Wall and J. K. Pritchard. Haplotype blocks and linkage disequilibrium in the human genome. *Nat Rev Genet*, 4(8):587–97, 2003.

[545] P J Waller, G Bernes, S M Thamsborg, A Sukura, S H Richter, K Ingebrigtsen, and J Höglund. Plants as de-worming agents of livestock in the nordic countries: historical perspective, popular beliefs and prospects for the future. *Acta Veterinaria Scandinavica*, 42(1):31–44, 2001. PMID: 11455900.

[546] Peter J Waller and Stig M Thamsborg. Nematode control in 'green' ruminant production systems. *Trends in parasitology*, 20(10):493–497, October 2004. PMID: 15363444.

[547] G. A. Walling, C. S. Haley, M. Perez-Enciso, R. Thompson, and P. M. Visscher. On the mapping of quantitative trait loci at marker and non-marker locations. *Genet Res*, 79(1):97–106, 2002.

[548] Zhong Wang, Mark Gerstein, and Michael Snyder. RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews. Genetics*, 10(1):57–63, January 2009. PMID: 19015660.

[549] S W Wanyangu, J M Mugambi, R K Bain, J L Duncan, M Murray, and M J Stear. Response to artificial and subsequent natural infection with haemonchus contortus in red maasai and dorper ewes. *Veterinary Parasitology*, 69(3-4):275–282, May 1997. PMID: 9195737.

[550] B. S. Weir and W. G. Hill. Effect of mating structure on variation in linkage disequilibrium. *Genetics*, 95(2):477–88, 1980.

[551] J I Weller, Y Kashi, and M Soller. Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. *Journal of dairy science*, 73(9):2525–2537, September 1990. PMID: 2258496.

[552] J I Weller and M Ron. Invited review: quantitative trait nucleotide determination in the era of genomic selection. *Journal of dairy science*, 94(3):1082–1090, March 2011. PMID: 21338774.

[553] J C Whittaker, R Thompson, and M C Denham. Marker-assisted selection using ridge regression. *Genetical research*, 75(2):249–252, April 2000. PMID: 10816982.

[554] P. Wiener and S. Wilkinson. Deciphering the genetic basis of animal domestication. *Proceedings of the Royal Society B: Biological Sciences*, 278(1722):3161–3170, September 2011.

[555] G R Wiggans, T S Sonstegard, P M VanRaden, L K Matukumalli, R D Schnabel, J F Taylor, F S Schenkel, and C P Van Tassell. Selection of single-nucleotide polymorphisms and quality of genotypes used in genomic evaluation of dairy cattle in the united states and canada. *Journal of dairy science*, 92(7):3431–3436, July 2009. PMID: 19528621.

[556] Valerie M Williamson and Amar Kumar. Nematode resistance in plants: the battle underground. *Trends in genetics: TIG*, 22(7):396–403, July 2006. PMID: 16723170.

[557] Perrin M Wilson, Robert H Fryer, Yin Fang, and Mary E Hatten. Astn2, a novel member of the astrotactin gene family, regulates the trafficking of ASTN1 during glial-guided neuronal migration. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 30(25):8529–8540, June 2010. PMID: 20573900.

[558] R G Windon. Selective breeding for the control of nematodiasis in sheep. *Revue scientifique et technique (International Office of Epizootics)*, 9(2):555–576, June 1990. PMID: 2132696.

[559] Virginia D Winn, Matthew Gormley, Agnes C Paquet, Kasper Kjaer-Sorensen, Anita Kramer, Kristen K Rumer, Ronit Haimov-Kochman, Ru-Fang Yeh, Michael T Overgaard, Ajit Varki, Claus Oxvig, and Susan J Fisher. Severe preeclampsia-related changes in gene expression at the maternal-fetal interface include sialic acid-binding immunoglobulin-like lectin-6 and pappalysin-2. *Endocrinology*, 150(1):452–462, January 2009. PMID: 18818296.

[560] Mike D Winter. Nematodirus battus 50 years on–a realistic vaccine candidate? *Trends in Parasitology*, 18(7):298–301, July 2002. PMID: 12379949.

[561] C T Wittwer, M G Herrmann, A A Moss, and R P Rasmussen. Continuous fluorescence monitoring of rapid cycle DNA amplification. *BioTechniques*, 22(1):130–131, 134–138, January 1997. PMID: 8994660.

[562] Wojciech Wojciechowski, David P Harris, Frank Sprague, Betty Mousseau, Melissa Makris, Kim Kusser, Tasuko Honjo, Katja Mohrs, Markus Mohrs, Troy Randall, and Frances E Lund. Cytokine-producing effector b cells regulate type 2 immunity to h. polygyrus. *Immunity*, 30(3):421–433, March 2009. PMID: 19249230.

[563] B.T. Wolf, K. Howells, C. Nakielny, W. Haresign, R.M. Lewis, O. Davies, and M.H. Davies. Genetic parameters for strongyle and nematodirus faecal egg counts in lambs and their relationships with performance traits. *Livestock Science*, 113(2-3):209–217, February 2008.

[564] Adrian J. Wolstenholme, Ian Fairweather, Roger Prichard, Georg von Samson-Himmelstjerna, and Nicholas C. Sangster. Drug resistance in veterinary helminths. *Trends in Parasitology*, 20(10):469–476, October 2004.

[565] R R Woolaston and R L Baker. Prospects of breeding small ruminants for resistance to internal parasites. *International Journal for Parasitology*, 26(8-9):845–855, September 1996. PMID: 8923134.

[566] R R Woolaston and R.G. Windon. Selection of sheep for response to trichostrongylus colubriformis larvae: genetic parameters. *Animal Science*, 73:41–48, 2001.

[567] R.R. Woolaston, R.L. Elwin, and I.A. Barger. No adaptation of haemonchus contortus to genetically resistant sheep. *International Journal for Parasitology*, 22(3):377–380, May 1992.

[568] Mark E J Woolhouse, Joanne P Webster, Esteban Domingo, Brian Charlesworth, and Bruce R Levin. Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nature genetics*, 32(4):569–577, December 2002. PMID: 12457190.

[569] Sitao Wu, Robert W Li, Weizhong Li, Ethiopia Beshah, Harry D Dawson, and Jr Urban, Joseph F. Worm burden-dependent disruption of the porcine colon microbiota by trichuris suis infection. *PloS one*, 7(4):e35470, 2012. PMID: 22532855.

[570] Shizhong Xu. Estimating polygenic effects using markers of the entire genome. *Genetics*, 163(2):789–801, February 2003. PMID: 12618414.

[571] T. A. Yazwinski, L. Goode, D. J. Moncol, G. W. Morgan, and A. C. Linnerud. Haemonchus contortus resistance in straightbred and crossbred barbados blackbelly sheep. *J Anim Sci*, 51(2):279–84, 1980.

[572] Anna R Young, Garry J Barcham, Joanna M Kemp, Jillian L Dunphy, Andrew Nash, and Els N Meeusen. Functional characterization of an eosinophil-specific galectin, ovine galectin-14. *Glycoconjugate journal*, 26(4):423–432, May 2009. PMID: 18810635.

[573] F. Ytournel. *Déséquilibre de liaison et cartographie de QTL en population sélectionnée.* PhD thesis, Agroparistech, Paris, 2008.

[574] Cheng Yue, Gerald Coles, and Neil Blake. Multiresistant nematodes on a devon farm. *The Veterinary Record*, 153(19):604, November 2003. PMID: 14640331.

[575] G. U. Yule. On the association of attributes in statistics. *Philos. Trans. R. Soc. London A.*, 194:257–139, 1900.

[576] Z. B. Zeng. Precision mapping of quantitative trait loci. *Genetics*, 136(4):1457–68, 1994.

[577] Bing Zhang, Stefan Kirov, and Jay Snoddy. WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic acids research*, 33(Web Server issue):W741–748, July 2005. PMID: 15980575.

[578] Shuanglin Zhang, Xiaofeng Zhu, and Hongyu Zhao. On a semiparametric test to detect associations between quantitative traits and candidate genes using unrelated individuals. *Genetic epidemiology*, 24(1):44–56, January 2003. PMID: 12508255.

[579] H. Zhao, D. Nettleton, M. Soller, and J. C. Dekkers. Evaluation of linkage disequilibrium measures between multi-allelic markers as predictors of linkage disequilibrium between markers and QTL. *Genet Res*, 86(1):77–87, 2005.

[580] H H Zhao, R L Fernando, and J C M Dekkers. Power and precision of alternate methods for linkage disequilibrium mapping of quantitative trait loci. *Genetics*, 175(4):1975–1986, April 2007. PMID: 17277369.

[581] M. Zuker. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*, 31(13):3406–3415, 2003.

AUTHOR: Guillaume SALLÉ

TITLE: Detection and functional validation of genomic regions affecting resistance to gastro-intestinal nematodes in sheep

---

SUMMARY:

Gastro-intestinal nematodes, among which *Haemonchus contortus* are a major threat to the meat sheep industry. They are responsible for production losses and the apparition of worm populations resistant to drugs limits their use as worm control strategy. Breeding more resistant sheep is among the most practicable alternative strategy. However its implementation requires a deeper understanding of underlying mechanisms. This PhD aims at identifying regions of the ovine genome affecting resistance to gastro-intestinal nematodes.

A statistical analysis of existing associations between genetic markers and resistance traits of a Martinik Black-belly x Romane cross-bred sheep flock unraveled a limited number of key players. Among these, a fragment of the chromosome 12 was chosen to perform marker-assisted matings and to validate its role in resistance to *H. contortus*. The effect of this region was validated in the progenies born from matings. It seems this chromosomic fragment limits female worms fertility and is associated to a stronger immune response.

The putative role played by a fragment of the chromosome 21 in plasma pepsinogen concentration (a biomarker of abomasal lesions) was also confirmed in this work. A candidate gene underlying this region has been sequenced and the analysis of the detected polymorphisms should confirm its role. Further, two other genes in its vicinity could also play a role in this biological phenomenon and they should also deserve future considerations.

This work illustrated both the existing genetic variation for resistance to *H. contortus* and the associated complexity of underlying mechanisms. Additional sequencing and gene expression sequencing studies should help understanding gene functions and interactions.

AUTEUR: Guillaume SALLÉ

TITRE: Détection et validation fonctionnelle de régions du génome affectant la résistance aux strongles gastro-intestinaux chez le mouton

DIRECTEURS DE THESE: Philippe JACQUIET et Carole MORENO-ROMIEUX

LIEU ET DATE DE SOUTENANCE: TOULOUSE, le 17 Décembre 2012

RESUMÉ EN FRANCAIS:

Les strongles gastro-intestinaux, dont *Haemonchus contortus* constituent un problème majeur pour l'élevage des ovins allaitants. Ils entrainent des pertes de production et le recours aux anthelminthiques est remis en question par l'apparition de souches de vers résistantes. La sélection d'ovins plus résistants fait partie des stratégies complémentaires de lutte les plus sérieuses. Cependant sa mise en oeuvre requiert une meilleure compréhension des mécanismes sous-jacents. Cette thèse vise à identifier les régions du génome ovin impliquées dans la résistance aux strongles gastro-intestinaux.

Une analyse statistique d'association entre des marqueurs génétiques et des mesures de résistance d'un troupeau d'ovins croisés Martinik Black-belly x Romane a mis en évidence un nombre limité de régions d'intérêt. Parmi celles-ci, un segment du chromosome 12 a été choisi pour effectuer des accouplements raisonnés et valider son rôle dans la résistance à *H. contortus*. L'effet de cette région a été validé chez les descendants issus d'accouplements assistés par marqueurs génétiques. Cette région semble limiter fertilité des vers femelles tout en contribuant à une réponse immunitaire plus forte.

Le rôle d'une région du chromosome 21 dans la variation de concentration plasmatique en pepsinogène, un marqueur de lésions abomasales, a également été confirmé. Un gène candidat sous-jacent est en cours de séquençage et l'analyse des polymorphismes devrait contribuer à la validation de son rôle. Deux autres gènes très proches pourraient également être impliqués et mériteraient une considération future.

Ces travaux illustrent à la fois la variation génétique disponible pour les caractères de résistance à *H. contortus* et la complexité des mécanismes mis en jeu. Des études complémentaires de séquençage et d'étude d'expression par séquençage devrait contribuer à une meilleure compréhension des fonctions des gènes impliqués et de leurs interactions.

MOTS-CLÉS: QTL, nématode, résistance génétique, SNP, expression de gènes, pepsinogène, analyse de liaison, analyse d'association, LDLA, déséquilibre de liaison

DISCIPLINE: Génétique animale et parasitologie

LABORATOIRES D'ACCUEIL:

- INRA/INP-ENVT, UMR1225, Interactions Hôtes-Agents Pathogènes, F-31076 Toulouse, France
- INRA, UR631, Station d'Amélioration Génétique des Animaux, F-31320 Castanet-Tolosan, France